52

Identification and Characterization of Novel *Rhizobium meliloti* Genes Involved in Carbon Metabolism

by

Laura B. Willis

B. S. *summa cum laude* Genetics and Cell Biology
University of Minnesota, 1990

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN BIOLOGY
AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 1998

Signature of Author: .....
Department of Biology
January 28, 1998

Certified by: .....
Graham C. Walker
Professor of Biology
Thesis Supervisor

Accepted by: .....
Frank Solomon
Professor of Biology
Chairman, Committee for Graduate Students

TABLE OF CONTENTS

Identification and Characterization of Novel *Rhizobium meliloti* Genes Involved in Carbon Metabolism

by

Laura B. Willis

Submitted to the Department of Biology
on January 28, 1998 in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy in Biology

ABSTRACT

During the symbiosis between the gram-negative soil bacterium *Rhizobium* (*Sinorhizobium*) *meliloti* strain and the leguminous plant *Medicago sativa* (alfalfa), the *R. meliloti* fix atmospheric nitrogen into ammonia which can be used by the plant, and the plant provides fixed carbon. The carbon storage compound poly-beta-hydroxybutyrate (PHB) is produced by *R. meliloti* in the free-living state and during invasion but disappears after the bacteria have invaded the plant and before nitrogen fixation begins. We cloned and sequenced the *R. meliloti phbC* gene encoding PHB synthase, constructed *R. meliloti phbC* mutants and showed that they are unable to produce PHB but are still able to elicit nitrogen-fixing nodules. The major form of carbon which is transported to the root nodules of leguminous plants is sucrose. Since we were interested in sucrose metabolism, we isolated *R. meliloti* genomic clones which allow a heterologous bacterial strain to utilize sucrose. A *R. meliloti* alcohol dehydrogenase gene (*adhA*) was found to permit growth of an *Alcaligenes eutrophus phbC* mutant on nitrogen limiting minimal medium containing sucrose. *adhA* is expressed in *A. eutrophus* under conditions where the native alcohol dehydrogenase is not expressed. A model for the role of *adhA* is presented. A second screen for genes involved in sucrose uptake and hydrolysis led to the identification of 5 loci which allow *A. eutrophus* to utilize any of the alpha-glucosides sucrose, maltose or trehalose These loci evidently encode an alpha glucosidase (*aglA*) and a periplasmic binding-protein dependent sugar transport system (*aglE*, *aglF*, *aglG* and *aglK*). The *agl* genes permit uptake of radiolabeled sucrose into *A. eutrophus*. *R. meliloti aglA* mutants grow well on alpha-glucosides, suggesting that *R. meliloti* possesses at least one additional alpha-glucosidase. *R. meliloti aglE*, *aglF* and *aglG* mutants grow poorly on alpha-glucosides, suggesting that *aglEFGK* encode the primary transport system for these sugars. By mutagenizing *agl* mutant strains it may be possible to isolate double mutants unable to utilize alpha-glucosides. Such mutants would make it possible to evaluate the importance of sucrose metabolism during symbiosis.

Thesis Supervisor: Graham C. Walker

Title: Professor of Biology

For my dearest friends and family

Acknowledgments

Many thanks to my advisor, Dr. Graham C. Walker, who has struck a fine balance between mentor and colleague. His encouragement and enthusiasm have been invaluable. A special thank you to the members of my thesis committee, Drs. Ethan Signer, Phillips Robbins, Anthony Sinskey, Andrew Wright and Gene Brown, for the interest they have shown in my work and my scientific development.

Heartfelt thanks go to all the members of the Walker lab for providing a fun and stimulating place to work. Thanks to Asao Ichige for his wisdom and kindness, Tonya Anderson for her friendship and cheerful whistling, Sumati Murli for her biting wit, Tom Barthel for the airplane ride, Brett Pellock for being so darned helpful, Brad Smith for his kvetching and listening, Kristin LeVier for embodying the light at the end of the tunnel and Greg York for keeping a handle on things. I gratefully acknowledge the many colleagues who have made Cambridge a real scientific community, especially Oliver Peoples and Tillmann Gerngross who provided strains and assistance with *Alcaligenes eutrophus*. Thanks to Dan Gage for providing unpublished results. This work was supported in part by a National Science Foundation predoctoral fellowship and supplemental funding from the Howard Hughes Medical Institute, with additional support from a National Institutes of Health grant to Dr. Graham C. Walker.

This thesis would not have been possible without the support and encouragement of my family. Thanks to my parents, artists and scientists both, for nurturing those talents in their children, and to my brother for his precious and timely reality checks. Thanks to Jaine for being there when I needed her most. And especially to David, my dearest friend and confidant.

Chapter 1

Introduction

All living things require the element nitrogen as an essential building block for cellular components including proteins and nucleic acids, but most are unable to utilize the essentially inert N_2 gas which comprises 78% of the earth's atmosphere. Instead, they rely on specialized bacteria and cyanobacteria which are able to convert nitrogen gas into ammonia in a process called biological nitrogen fixation. Most nitrogen fixers are soil or water organisms, and some can only fix nitrogen when they are in associative or symbiotic relationships with plants or fungi. One well characterized group of symbiotic nitrogen fixing bacteria is the Rhizobiaciae, which includes the genera *Rhizobium*, *Bradyrhizobium*, *Sinorhizobium* and *Azorhizobium*. The Rhizobiaciae fix nitrogen inside specialized structures known as nodules on the roots of plants of the family Leguminosae.

Biological nitrogen fixation is a process of great agroeconomic importance. Many familiar food crops are legumes, including alfalfa, soybean, and many of types of peas and beans. Other plants which are remarkable for their fast growth, such as kudzu and black locust, are also legumes. The well established practice of crop rotation involves planting legumes and non-legumes in the same field in alternate years. The legume-associated rhizobia enrich the soil with fixed nitrogen that can be utilized by the following year's planting. By studying the symbioses between these plant hosts and their bacterial guests, scientists have learned a great deal about both the bacterial machinery which carries out the nitrogen fixation and the complicated and elegantly balanced communication that goes on between the macrosymbiont and microsymbiont partners.

This work focuses on the interaction between one of these pairs of partners: the gram-negative soil bacterium *Rhizobium meliloti* and the leguminous plant *Medicago sativa* (alfalfa). During the development of the *R. meliloti-M. sativa* symbiosis, both partners undergo changes in gene expression which lead to distinct

morphological changes. The process begins when chemical signaling between the symbionts initiates a developmental pathway in the plant, resulting in the formation of a new organ on the root of the plants: the nodule (reviewed in Fisher and Long, 1992; van Rhijn and Vanderleyden, 1995).

Chitooligosaccharide Nod factors

Alfalfa plants secrete flavonoids into the rhizosphere, which act as rhizobial chemoattractants (Caetano-Anolles et al., 1988) and inducers of bacterial *nod* gene expression (Peters et al., 1986; Redmond et al., 1986). Some of the most important advances in the *Rhizobium* field in recent years have been the elucidation of *nod* gene function (Denarie and Cullimore, 1993; Vijn et al., 1993) and the very early steps of nodule initiation. The *nod* gene products are required for the synthesis of Nod factors, molecules with a backbone of 3 to 5 residues of N-acetyl glucosamine modified with an N-acyl lipid moiety at the non-reducing end (Denarie and Cullimore, 1993). Different species of rhizobia produce distinct Nod factors (Spaink et al., 1991; Sanjuan et al., 1992; Mergaert et al., 1993). Rhizobial host range is determined in part by which plants respond to a particular Nod factor (Schwedock and Long, 1990; Spaink et al., 1995). Nod factor is able to induce root hair curling and cell division when added to seedlings in the absence of bacteria (Lerouge et al., 1990; Truchet et al., 1991). In addition, purified *Rhizobium meliloti* Nod factor has been shown to depolarize the alfalfa root hair cells (Erhardt et al., 1992). This plant response may be responsible for the changes in gene expression needed for nodule development.

One unexpected finding of these studies was the discovery that chitooligosaccharide molecules have biological activity in vastly different systems. For example, exogenous application of purified Nod factor was found to rescue the

developmental block in a carrot embryogenesis mutant (De Jong et al., 1993), suggesting that this non-legume could use a similar molecule as an endogenous signal. The *Xenopus* DG42 gene, which is expressed only during embryogenesis, was found to encode a protein capable of synthesis of Nod-factor-like chitooligosaccharides (Semino and Robbins, 1995), and homologues of DG42 have been cloned from zebrafish and mouse (Semino et al., 1996). In addition, Bakkers et al. have shown that chitooligosaccharide production is required during a defined stage of embryogenesis in zebrafish (Bakkers et al., 1997). These results demonstrate that chitooligosaccharides define a new class of developmental signals found in both plants and animals.

Alfalfa root nodule development

After the initiation of alfalfa nodule development has been stimulated by Nod factors, cortical cell divisions begin and the nodule primordium grows outward from the root (reviewed in Brewin, 1991; Franssen et al., 1992; Hirsch, 1992). *R. meliloti* induces curling of alfalfa root hair cells, and then colonizes and invades these cells. The bacteria physically invade the plant through a tube which forms inside the root hair. This tube, termed the infection thread, grows from the point of colonization inward toward the dividing plant cells of the nodule primordium. After traveling through the infection thread, the invading bacteria are engulfed (in a process resembling endocytosis) within a membrane of plant origin, forming the symbiosome organelle. Because they are encased in this symbiosome membrane, the bacteria remain topologically outside of the plant cell.

After they are released into this new environment, the bacteria begin to differentiate into a larger and sometimes branched form called the bacteroid. Bacteroid development proceeds through a series of four well defined

developmental stages (Vasse et al., 1990). Mature bacteroids are able to fix atmospheric nitrogen to ammonia, which can be transported and utilized by the plant. Throughout the processes of invasion and bacteroid maturation, the plant provides fixed carbon to the bacteria, as it will continue to do during the course of the interaction.

Bacterial mutants have been very useful for genetic dissection of the nodule developmental pathway. Many *R. meliloti* genes required for the early steps of nodule development (Debelle et al., 1986; Honma et al., 1990) and invasion (Leigh et al., 1985; Reed and Walker, 1991a; Reed and Walker, 1991b; Glucksmann et al., 1993a; Glucksmann et al., 1993b) as well as the later stage of nitrogen fixation (Putnoky et al., 1988; Kahn et al., 1993) have been well characterized. Analysis of *R. meliloti* strain Rm1021 has shown that bacteria which fail to produce the acidic exopolysaccharide succinoglycan are not able to invade nodules (Leigh et al., 1985; Long et al., 1988; Leigh and Coplin, 1992), and elicit the formation of white, ineffective nodules which contain few or no bacteria and no bacteroids. Production of a normally cryptic polysaccharide called EPS II can rescue this defect (Glazebrook and Walker, 1989), and exogenously applied low molecular weight EPS II is able to rescue a *R. meliloti* strain that produces neither exopolysaccharide, leading to the production of effective nodules (González et al., 1996).

A class of plant host genes, termed nodulins, which are expressed exclusively or differentially in nodules, has been identified by analyzing mRNAs isolated from roots or nodules (Legocki and Verma, 1980; Lullien et al., 1987). One class of nodulins is comprised of oxygen-binding proteins known as leghemoglobins (Verma et al., 1981) which are responsible for the characteristic pink color of effective nodules. These proteins bind a heme group which is produced by the bacteroids (O'Brian, 1996) and sequester most of the oxygen in the nodule, creating

the low oxygen conditions necessary for activity of the nitrogenase complex. By examining which nodulins are expressed in spontaneous nodules in comparison to nodules elicited by bacterial mutants it has been possible to show that expression of the late nodulins, such as the leghemoglobins, may be dependent on release of bacteroids into the symbiosome membrane but is not dependent on the ability of the bacteroids to fix nitrogen (Lullien et al., 1987; Norris et al., 1988).

Carbon storage compounds observed in alfalfa nodules

Some nodules, such as those elicited on soybean roots by *Bradyrhizobium japonicum*, grow to a certain size and then stop. They are therefore categorized, using plant biology terminology, as determinate. In contrast, the nodules elicited on alfalfa roots by *R. meliloti* continue to grow outward from the root for several weeks. Nodules of this type are classified as indeterminate. A longitudinal section through an indeterminate nodule can be interpreted as a timeline of the plant-microbe interaction (Hirsch, 1992). The most plant-distal point of this type of nodule is the meristem, where active plant cell division is taking place. Immediately proximal to the meristem is the invasion zone, where *R. meliloti* invade some of the plant cells. Separating the invasion zone from the older regions of the nodule is a band of plant cells 2 to 3 cells thick (Vasse et al., 1990), in which the uninfected plant cells contain numerous starch granules. These prominent starch granules are not seen in the mature nitrogen fixing zone of the nodules (Vasse et al., 1990). Starch deposition is increased in ineffective and spontaneous nodules (Joshi et al., 1991), a striking example of which is seen in the nodules elicited by *exo* mutants of *R. meliloti* (Finan et al., 1985; Leigh et al., 1987). It appears that the nodule developmental pathway in the plant directs the production and storage of starch, which disappears when effective symbiosis is achieved.

In an analogous situation in the bacteria, *R. meliloti* in the infection thread contains numerous granules of the electron-transparent carbon storage compound poly-β-hydroxybutyrate (Paau et al., 1978). However, these granules disappear after the bacteria are released inside the symbiosome membrane, and this disappearance marks the first stage in bacteroid development (Vasse et al., 1990). It appears that both the plant and the rhizobia are storing carbon just prior to the establishment of an effective symbiosis. In Chapter 2 we describe the isolation and characterization of *R. meliloti* mutants unable to synthesize poly-β-hydroxybutyrate.

Carbon metabolism by *R. meliloti*

Studies of carbon metabolism by *R. meliloti* are complicated by the distinction between carbon sources used in the free-living state and those used *in planta*. Several groups have studied sugar metabolism by free-living rhizobia. The lactose utilization operon of *R. meliloti* has been cloned and sequenced and consists of a β-galactosidase (*lacZ*), a permease (*lacY*) and two other loci of unknown function (*lacW* and *lacX*) (Jelesko and Leigh, 1994). *R. trifolii* mutants defective in disaccharide metabolism have been characterized and are not impaired in their ability to induce effective nodules (Ronson and Primrose, 1979). *R. meliloti* mutants defective in galactose metabolism (Arias and Cerveñansky, 1986) have also been characterized. The carbon source(s) utilized by bacteria in the infection thread and developing bacteroids have not been elucidated, although it has been shown that *R. meliloti* genes involved in catabolism of the plant secondary metabolite trigonelline are induced during invasion and *in planta* (Boivin et al., 1990).

Much of the research on carbon metabolism in the *R. meliloti-M. sativa* interaction has focused on genes that are required to fulfill the high energy requirements for activity of the nitrogenase holoenzyme. It has been shown

conclusively that transport of dicarboxylic acids into the bacteroid is required for effective symbiosis (Ronson et al., 1981; Bolton et al., 1986; Engelke et al., 1987; Yarosch et al., 1989; Jording et al., 1994). *dct* mutants cannot transport dicarboxylic acids, rendering them unable to grow on carbon sources such as succinate in the free living state (Finan et al., 1988). In the symbiotic interaction, failure to transport succinate into the bacteroid reduces the activity of the tricarboxylic acid (TCA) cycle. *dct* mutants are Fix^-, and it has been proposed that they do not produce enough energy *in planta* to permit the activity of nitrogenase. Consistent with the model that an active TCA cycle is required for biological nitrogen fixation, *Rhizobium* mutants which lack the activity of TCA cycle enzymes are Fix^- (Driscoll and Finan, 1993). Although the nodules they induce are ineffective, *dct* mutants are still able to induce and invade nodules and proceed through the first three stages of bacteroid development (Vasse et al., 1990) despite their inability to utilize dicarboxylic acids as a carbon source. Therefore, there must exist additional pathways for carbon transport and utilization during invasion and subsequent bacteroid development.

R. meliloti sucrose metabolism

The disaccharide sucrose is readily utilized by free living *R. meliloti*, and is found in both the plant and bacteroid fractions of alfalfa root nodules. When nodulated leguminous plants are provided $^{14}CO_2$, the first radiolabeled sugar detected in nodule and bacteroid fractions is sucrose (Romanov et al., 1985; Streeter, 1991). As a non-reducing sugar, sucrose is less susceptible to modification and thus is the main sugar transported in the phloem (Giaquinta, 1983). Its presence at high concentrations in the bacteroids suggests that sucrose is being transported across the symbiosome membrane. Transport of sucrose in the infection thread has not been investigated.

No genes involved in *R. meliloti* sucrose metabolism have been identified, nor have mutants of *R. meliloti* have been reported that were isolated on the basis of their inability to transport or cleave sucrose. No evidence of sucrose phosphorylase activity, required for sucrose uptake via the PEP-phosphotransferase system (PTS) utilized by enteric bacteria, has been found in fast- or slow-growing rhizobia (Martinez-de Drets et al., 1974). *R. meliloti* has been shown to produce both sucrose uptake and hydrolysis activities (Martinez-de Drets et al., 1974; Glenn and Dilworth, 1981). Disaccharide uptake in *R. meliloti* is an active process, and competition studies suggest that the α-glucoside disaccharides sucrose, trehalose and maltose are imported by the same transport system (Glenn and Dilworth, 1981).

Sucrose-cleaving enzymes

As the main photosynthate transported in the phloem of higher plants (Giaquinta, 1983), sucrose is very abundant and many organisms have evolved ways to utilize this disaccharide. By convention, the glycosidic bond in sucrose is classified as an α-glucosidic bond. But viewed from the perspective of the fructose, the bond is a β-linkage. For this reason, both α-glucosidases and β-fructofuranosidases have been identified with sucrose-cleaving activity.

Hundreds of genes encoding glycosyl hydrolases have been cloned and sequenced, providing a large body of information about these enzymes. Glycosyl hydrolases have been divided into families based on substrate specificity and protein sequence (Henrissat, 1991; Henrissat and Bairoch, 1993; Henrissat and Romeu, 1995). Many sucrose cleaving enzymes belong to family 13 of glycosyl hydrolases (Janecek et al., 1997), which is also called the α-amylase family to reflect the fact that many of its members are active on amylose (starch). This family contains proteins with at least 18 different substrate specificities (Svensson, 1994), including α-amylase,

α-glucosidase (Suzuki et al., 1989), oligo-1,6-glucosidase (Kizaki et al., 1993), and starch branching enzyme (Poulsen and Kreiberg, 1993). Other enzymes that cleave sucrose include extracellular and intracellular invertases, sucrose-6-phosphate hydrolase (Schmid et al., 1982), and levansucrase (Cruz et al., 1990), which cleaves sucrose and catalyzes the formation of long polymers of fructose subunits.

Sucrose metabolism by other bacteria

Although *Escherichia coli* K12 is unable to utilize sucrose, sucrose positive *E. coli* are found in nature (Palchaudhuri et al., 1977; Lengeler et al., 1982). One isolate, EC3132, was found to encode a sucrose hydrolase (CscA), D-fructokinase (CscK) and a proton symport type transport system (CscB) (Bockman et al., 1992). The *csc* locus is an inducible regulon and also encodes a sucrose-specific repressor (CscR). Other sucrose positive *E. coli* carry the genes necessary for sucrose metabolism on plasmids such as pUR400 (Schmid et al., 1982), which encodes a sucrose transporter (*scrA*) and sucrose-6-phosphate hydrolase (*scrB*) and relies on the host's intact phospho-enolpyruvate:carbohydrate phosphotransferase system (PTS).

The PTS couples the phosphorylation of substrates with their transport across the inner membrane (Postma et al., 1993; Saier and Reizer, 1994). The PTS is also involved in chemotaxis and regulation of the synthesis and activity of other transport systems. This is a modular system, with two general enzymes and one or two substrate-specific enzymes which interact with the general enzymes. Enzyme I acquires a phosphate from phosphoenolpyruvate and transfers it to HPr. Enzyme II is a membrane protein that transports a specific substrate, which is then phosphorylated by HPr-P. In some cases a fourth protein, Enzyme III, is associated with Enzyme II. The genes encoding Enzyme I and HPr, *ptsI* and *ptsH*, respectively, are linked on the *E. coli* chromosome. The *scrA* gene from pUR600 (Schmid et al., 1982) encodes a sucrose-specific Enzyme II (Titgemeyer et al., 1996).

The majority of sucrose-positive bacteria which have been reported use the PTS for transport and hydrolysis of sucrose (Postma and Lengeler, 1985; Reizer et al., 1988). A well-characterized example is the gram-positive bacterium *Bacillus subtilis*. Expression of the *B. subtilis sacPA* operon, encoding a sucrose-specific PTS-dependent permease (SacP) and an intracellular sucrase (SacA), is regulated by SacT (Debarbouille et al., 1990). The expression of *sacB*, encoding levansucrase, is negatively regulated by SacX, which shares homology with SacP, and positively regulated by SacY (Cruz et al., 1990). SacT and SacY are transcriptional antiterminators which, in the presence of sucrose, bind to the nascent RNA and permit transcription to proceed past ρ-independent terminators found upstream of the *sacPA* and *sacB* genes (Arnaud et al., 1996). The regulation of *sacPA* and *sacB* by antitermination strongly resembles that observed at the *bgl* operon of *E. coli*, where, in the presence of β-glucosides, BglG acts as an antiterminator to permit transcription of *bglF* and *bglB* (Houman et al., 1990).

Identification of sucrose utilization genes

Many systems for the utilization of sugars have been identified by isolation of clones which allow *E. coli* K12 to utilize sucrose. For example, the *Staphylococcus xylosus scrB* gene, which encodes a sucrase, was isolated by screening a *S. xylosus* genomic library in *E. coli* (Brueckner et al., 1993). An operon encoding a regulatory protein (MalR) and an α-1,4-glucosidase (MalA) involved in maltose-maltotriose utilization was isolated from the same organism using a similar strategy (Egeter and Brueckner, 1995). The *sacC* gene, encoding an extracellular sucrase, has been cloned from *Zymomonas mobilis* and expressed in *E. coli* (Kannan et al., 1995). An extracellular sucrase (*scrB*) has also been characterized from *Vibrio alginolyticus* and when expressed in *E. coli* the ScrB protein is translocated to the periplasm but not exported across the cytoplasmic membrane (Scholle et al., 1989).

Screening for sucrose metabolism genes in a heterologous host can yield unexpected results. For example, Schuerman et al. found a fragment of *Agrobacterium tumefaciens* DNA which gives *E. coli* the ability to grow on sucrose (Schuerman et al., 1997). However, when they examined the enzyme extracts of the sucrose-positive *E. coli*, instead of detecting invertase activity they found evidence of D-glucoside-3-dehydrogenase and α-3-ketoglucosidase activities. Further tests revealed that the latter enzyme was active on 3-ketosucrose, and the researchers suggest that the *A. tumefaciens* DNA fragment also confers the ability to modify sucrose so that it can be cleaved by the ketoglucosidase. In an unexpected result described in Chapter 3, an alcohol dehydrogenase gene was identified in a screen for *R. meliloti* genomic clones which permit the growth of an *A. eutrophus phbC* mutant on sucrose.

Chapter 4 describes the identification and characterization of *R. meliloti* genes involved in the metabolism of the α-glucosides sucrose, maltose and trehalose. These genes were found by screening a *R. meliloti* genomic library in *A. eutrophus*. *R. meliloti* appears to be the first bacterium reported to use a periplasmic binding protein dependent system that specifically transports sucrose, maltose and trehalose.

Chapter 2

The *phbC* (poly-β-hydroxybutyrate synthase) gene of *Rhizobium meliloti* and characterization of *phbC* mutants

ABSTRACT

Defined insertion mutations have been constructed in the *Rhizobium meliloti phbC* gene, which encodes poly-β-hydroxybutyrate (PHB) synthase. The locus was isolated and subcloned from a genomic library of *R. meliloti* Rm1021 by complementation of a *phbC* mutation of *Alcaligenes eutrophus*. PHB production was detected in wild type *R. meliloti* under limiting nutrient conditions, but not in rich medium. No PHB production was detected in the *R. meliloti phbC* mutants. The DNA sequence of the *R. meliloti phbC* gene was determined. The deduced polypeptide sequence is homologous to previously identified PhbCs from other bacteria. The *R. meliloti phbC* locus maps to pRmeSU47a, the smaller of two megaplasmids.

INTRODUCTION

Poly-β-hydroxybutyrate (PHB) is a carbon storage polymer produced by many bacteria, and a member of a class of biodegradable thermoplastics known as polyhydroxyalkanoates (PHAs) (Anderson and Dawes, 1990). Since its characterization by Lemoigne (Lemoigne, 1925; Lemoigne, 1926), PHB has been detected in many microorganisms. Early studies of PHB in the *Rhizobiaceae* focused on an apparent correlation between high levels of intracellular PHB and reduced efficiency of nitrogenase (Romanov et al., 1974; Tombolini and Nuti, 1989). Bacteroids within determinate nodules, such as those elicited on soybean by *Bradyrhizobium japonicum* and on bean by *Rhizobium etli*, often accumulate high levels of PHB which may be used as a carbon source during periods of darkness or seed formation, when photosynthate is unavailable (Bergersen et al., 1991) or as a sink for reducing equivalents.

In contrast, bacteroids within indeterminate nodules, such as those elicited on alfalfa by *Rhizobium meliloti*, do not accumulate PHB. During the invasion process *R. meliloti* traveling through the infection thread contain numerous granules of electron transparent poly-β-hydroxybutyrate (Paau et al., 1978). Upon release from the infection thread, the bacteria are enclosed in the plant-derived symbiosome membrane, and remain topologically outside of the plant. In response to the new environment of the symbiosome, the bacteria enter the bacteroid developmental pathway. Disappearance of the PHB granules is a hallmark of the first of four distinct stages of bacteroid development (Vasse et al., 1990).

Although the timing of this phenomenon is so precise and consistent it is used to define a stage of development, only one *R. meliloti* mutant has been reported which is blocked so early in bacteroid development that the PHB granules remain

within the cells. *R. meliloti bacA* mutants (Glazebrook et al., 1993) are able to initiate nodulation and invade growing nodules, but fail to differentiate into bacteroids. *bacA* cells senesce soon after engulfment in the symbiosome membrane, and the senescent and lysing *bacA* cells observed in electron micrographs contain PHB granules. BacA is a homologue of *Escherichia coli* SbmA, and both proteins appear to be membrane-spanning transporters. It is hypothesized that the early senescence is due to failure of the bacteroids to perceive a plant derived environmental signal, or alternately a failure to export a plant-derived toxin or bacterially-derived signaling molecule (Glazebrook et al., 1996; Ichige and Walker, 1997).

The proposed signal or molecule may act directly or indirectly as a regulator of gene expression. Some of the changes in gene expression during the transition from bacterium to bacteroid could include new transcription required for mobilization of the carbon in PHB granules. For example, genes encoding PHB depolymerase or other activities may be turned on in wild type cells during this transition, while genes encoding PHB biosynthetic enzymes may be turned off at this stage. In a *bacA* strain, these changes in gene expression do not occur presumably because the BacA protein is not present to transport the critical molecule(s), bacteroid development is blocked and the PHB granules are not broken down. By examining the production of PHB in the *R. meliloti-M. sativa* symbiosis we sought to better understand the role of PHB during invasion.

The PHB biosynthetic pathway has been elucidated in several organisms (Anderson and Dawes, 1990; Steinbüchel et al., 1992). Many bacteria, notably *Alcaligenes eutrophus*, produce PHB from acetyl-CoA in a three step process (Schubert et al., 1988; Peoples and Sinskey, 1989b; Peoples and Sinskey, 1989c) (Figure 2-1). In the first step of the pathway, biosynthetic β-ketothiolase condenses two

molecules of acetyl-CoA to form acetoacetyl-CoA. This substance is converted to β-hydroxybutyryl-CoA by a NADPH-dependent acetoacyl-CoA reductase. PHB synthase (often called PHA synthase or PHB polymerase) then catalyzes the polymerization of β-hydroxybutryl CoA to form PHB.

The PHB synthase (*phbC*) gene of *R. meliloti* strain 41 has been described (Tombolini et al., 1995). Rm41 has the same nodulation host range as Rm1021, but is a significantly different isolate. Clear distinctions can be made between the bacteriophage sensitivity profiles of Rm1021 and Rm41, reflecting differences in the cell surfaces of the two strains. Furthermore, Rm41 is capable of using a capsular polysaccharide, KPS, to invade nodules (Petrovics et al., 1993), while Rm1021 uses one of two exopolysaccharides to fulfill this function (Leigh et al., 1985; Glazebrook and Walker, 1989).

Rm1021 and its sibling strain Rm2011 have been used in a wide variety of studies of nodules (Hirsch et al., 1982; Dudley et al., 1987; Vasse et al., 1990; Glazebrook and Walker, 1991; Yang et al., 1992) and have been more extensively characterized than Rm41. We were interested in examining PHB synthesis in Rm1021 in order to gain insight into its possible influences on the nodulation process and to be better able to analyze the signals controlling PHB synthesis and degradation during symbiosis.

MATERIALS AND METHODS

Strains and growth media.

Bacterial strains and plasmids used in this study are listed in Table 2-1. *R. meliloti* was routinely grown in Luria-Bertani (LB) broth supplemented with 2.5 mM $MgSO_4$ and 2.5 mM $CaCl_2$. *A. eutrophus* was grown in LB broth or Trypticase Soy broth (Gibco/BRL). *E. coli* was grown in LB broth. The low nitrogen minimal medium MM1 (Peoples and Sinskey, 1989b), containing 0.5% fructose as the sole carbon source, was used to screen *A. eutrophus* strains for PHB production. Antibiotics were used at the following concentrations: ampicillin (Amp), 150 μg/ml; chloramphenicol (Cm), 20 μg/ml; gentamicin sulfate (Gm), 5 μg/ml for *E. coli*, 50 μg/ml for *R. meliloti*; kanamycin sulfate (Km), 50 μg/ml; nalidixic acid (Nal), 50 μg/ml; neomycin sulfate (Nm), 200 μg/ml; spectinomycin (Sp), 100 μg/ml; tetracycline (Tc), 10 μg/ml.

Genetic techniques

Conjugal transfer of plasmids was accomplished in triparental matings using pRK600 to provide transfer functions. In matings where *A. eutrophus* was the donor, a 5:1:1 ratio of donor:recipient:helper culture volumes was used. Generalized transduction in *R. meliloti* using bacteriophage ϕM12 and recombination of insertions into the *R. meliloti* genome were performed as described (Glazebrook and Walker, 1991).

DNA manipulations

Plasmid and cosmid DNA was isolated from overnight cultures of *E. coli* by the alkaline lysis method (Maniatis et al., 1982). *R. meliloti* and *A. tumefaciens* genomic DNA was prepared as described, using CTAB (hexadecyltrimethyl-

ammonium bromide) to precipitate contaminating polysaccharides (Ausubel et al., 1995). Restriction enzymes and DNA ligase were used according to the instructions of the supplier (New England Biolabs, Beverly, MA or Takara, Japan). Gene Screen Plus membranes (Dupont/NEN, Boston, MA) were employed for Southern blotting and hybridization. Hybridization probes were prepared with the NEBlot random labeling kit (New England Biolabs, Beverly, MA) and labeled with ^{32}P-α-dCTP from Dupont/NEN (Boston, MA).

DNA Sequencing and analysis

The 3.7 kb *Eco*RI fragment of pPHB6 was subcloned into pBluescript SK+ to form pLW113. Defined subclones of pLW113 were constructed using standard molecular biology techniques. The entire insert, 3.648 kb, was sequenced using Sequenase v. 2.0 (United States Biochemical), CircumVent PCR sequencing (New England Biolabs, Beverly, MA) or by the MIT Biopolymers Laboratory. Contigs were prepared with Assembly Lign software (Kodak/IBI) and the SeqMan program of the DNASTAR software package (Lasergene, Inc.). Database searches were performed using the BLAST (Altschul et al., 1990) algorithms and the databases maintained by the National Center for Biotechnology Information. Alignments and other analyses were performed using the GCG software package (Genetics Computer Group, 1991) and the DNASTAR software package. The DNA sequence reported in this chapter has been submitted to GenBank and assigned the accession number AF031938.

Isolation and analysis of complementing plasmids

A *R. meliloti* genomic library (Friedman et al., 1982) was mated into *A. eutrophus* PHB#2. Transconjugants were selected on MM1 supplemented with fructose, Tc and Km. After 3 days of incubation at 30° C, opaque, white PHB-producing strains

were isolated. The cosmids from these strains were mated into *E. coli* strain C2110 to facilitate isolation and analysis of cosmid DNA.

Construction of PHB$^+$ subclones

In separate reactions, pLW111-2 DNA was digested with *Eco*RI or *Bam*HI and then religated. These ligation mixes were used to transform *E. coli* DH5α competent cells and the resulting transformants were mated with *A. eutrophus* PHB#2. Transconjugants were selected on MM1 fructose Tc Km and opaque, white PHB-producing colonies were selected for further study. Cosmids from these strains were mated into *E. coli* C2110. The 3.7 kb insert from one PHB$^+$ subclone, pPHB6, was cloned into *Eco*RI-digested pSW213 to make pLW119.

Disruption of the R. meliloti phbC *gene*

pLW119 was digested with *Bam*HI and the 11 kb fragment containing the vector and most of the insert was gel purified from an agarose gel slice using the Qiaex extraction kit (Qiagen) and ligated with the kanamycin/neomycin resistance cassette of miniTn*5*Km (de Lorenzo et al., 1990) to construct pLW150. The kanamycin/neomycin resistance cassette was also cloned into the *Hin*dIII site of pLW119 to create pLW152. pLW150 and pLW152 were mated into Rm1021 and the drug resistance marker homogenotized into the genome. In each case the neomycin resistance marker was transduced into a fresh background of Rm1021 and the resultant strains Rm9600 and Rm9601 were used for further studies.

Mapping the phbC *locus*

The *phbC* locus was mapped to one of the *R. meliloti* replicons using the method of Finan et al. (Finan et al., 1985). Genomic DNA from Rm1021, Rm9600, Rm9601 and *Agrobacterium tumefaciens* strains At123, At125 and At128 was digested with *Eco*RI, electrophoretically separated on an 0.6% agarose gel and blotted onto a Gene Screen Plus filter (Dupont/NEN). The 1.2 kb *Bam*HI fragment from pLW113 was labeled with ^{32}P-α-dCTP and used to probe the filter.

Assays for PHB production

PHB assays were performed using the spectrophotometric technique described by Law and Slepeckey (Law and Slepeckey, 1961) as modified by Peoples and Sinskey (Peoples and Sinskey, 1989b). *A. eutrophus* or *R. meliloti* cultures were grown in LB medium or in MM1 fructose medium supplemented with the appropriate antibiotics. Pelleted cells from 50 ml cultures were resuspended in 1 ml lysis buffer [10 mM Tris·Cl pH 8.0; 5 mM β-mercaptoethanol; 5 mM EDTA (ethylenediamine tetraacetic acid); 0.02 mM phenylmethylsulfonyl fluoride (PMSF); 10% weight/volume glycerol], placed in an ice bath and sonicated. A microtip sonicator was used on setting 3, with pulses of 2.5 sec on, 2.5 sec off, for a total process time of 15 minutes.

The sonicate was divided in half, and one aliquot was used for isolation of PHB as follows. In a microcentrifuge tube, 100 µl of the sonicate (crude lysate) was combined with 1.2 ml of a 5% sodium hypochlorite solution and incubated at 37°C for 1 hour. The precipitated material was pelleted and washed with water, then actetone and finally ethanol. The pellet was dried under vacuum, then resuspended in 100 µl chloroform. 0.5, 5 and 50 µl aliquots were placed at the bottom of glass tubes, and the chloroform allowed to evaporate in the fume hood. The samples

were then dissolved in 5 ml of concentrated H_2SO_4 and placed in a boiling water bath (with glass marbles on top of the tubes) for 10 minutes. The OD_{235} was measured after the samples were allowed to cool. Purified PHB was purchased from Sigma (St. Louis, MO) for use as a concentration standard.

For determination of protein content, the second aliquot of sonicate was transferred to a microfuge tube and spun for 15 minutes in a tabletop microcentrifuge. The supernatant (cleared lysate) was used for determination of protein content by the method of Bradford (Bradford, 1976) using a commercially available reagent (Bio-Rad) and bovine serum albumin (New England Biolabs) as a standard.

Plant inoculation assays

Medicago sativa cv. Iroquois was obtained from Agway, Inc. (Plymouth, IN). *R. meliloti* strains were tested in alfalfa nodulation assays on nitrogen free Jensen's medium as described (Leigh et al., 1985). Alfalfa seedlings were inoculated with water, Rm1021, Rm7031 (an *exoA* derivative of Rm1021), Rm9600 or Rm9601. Plants were grown in a constant temperature room at 21° C with a 14 hour light cycle. Observations were made weekly for a minimum of six weeks. The presence of pink cylindrical nodules on healthy dark green plants was taken as evidence that nitrogen fixation was occurring. Plants lacking nodules or with ineffective nodules were stunted and chlorotic.

RESULTS

Isolation of R. meliloti *cosmids which complement an* A. eutrophus phbC *mutation*

To clone the *R. meliloti* strain Rm1021 *phbC* gene, we took advantage of a powerful screen for PHB production in *Alcaligenes eutrophus*. *A. eutrophus* will produce approximately 70-80% of its dry weight as PHB when grown under conditions of excess carbon and limitation for another nutrient (for example, nitrogen, sulfur or oxygen) (Anderson and Dawes, 1990; Schlegel and Steinbüchel, 1991). Under these conditions, colonies which produce PHB are opaque and white, whereas those which fail to produce PHB are translucent. This difference is clearly detected by eye.

We mated an *R. meliloti* genomic library into *A. eutrophus* strain PHB#2, which carries a Tn*5* insertion in the *phbC* gene encoding PHB synthase (Peoples and Sinskey, 1989b). Approximately 0.1% of the transconjugants produced opaque white colonies, and had the same colony morphology as wild type *A. eutrophus* grown on the same medium. This morphology suggested that the transconjugants are producing PHB. We selected twenty-five white colonies from five independent matings for further study.

We determined that the opaque white colony morphology was due to the presence of the cosmid by mating each cosmid into *E. coli* C2110 and back into a fresh background of *A. eutrophus phbC*. Each of the twenty-five cosmids conferred the opaque white colony morphology to *A. eutrophus phbC* grown on MM1 fructose, while an isogenic strain harboring pLAFR1 produced translucent colonies.

Isolation of PHB from A. eutrophus phbC *harboring* R. meliloti *genomic clones*

To ascertain whether the colony morphology was due to production of PHB or some other substance, we performed assays for PHB production. Results from a typical

assay are shown in Figure 2-2A. Cultures were grown to saturation in rich medium, then split for assay of protein and PHB. Material which is insoluble in bleach, water, acetone and ethanol but soluble in chloroform could be isolated from wild type *A. eutrophus* or *A. eutrophus phbC* strains harboring one of the *R. meliloti* genomic clones identified in this study, but not from *A. eutrophus phbC* harboring the vector pLAFR1. Moreover, this material has the same spectrophotometric properties as PHB, with a peak absorbance at 235 nm. Therefore, we concluded that the strains were producing PHB and that the *R. meliloti* genomic clones carried a gene or genes conferring PHB synthase activity. The levels of PHB production detected in *A. eutrophus phbC* harboring an *R. meliloti* genomic clone (0.8 mg PHB/mg protein) were lower than the levels of PHB produced by wild type *A. eutrophus* (2.5 mg PHB/mg protein).

R. meliloti *PHB Synthase activity is encoded on a 3.7 kb EcoRI fragment*

The 25 cosmids which confer PHB synthase activity on the *A. eutrophus* PHB$^-$ strain can be grouped into 8 distinct restriction patterns after digestion with *Eco*RI. The eight classes of cosmids appear to contain overlapping regions of the *R. meliloti* genome because many restriction fragments are present in several plasmids (Figure 2-3). One cosmid was chosen from each class; these eight cosmids were designated pLW111-1 through pLW111-8. The only *Eco*RI fragment present in each of the 25 cosmids was a 3.7 kb *Eco*RI fragment, which suggested that this fragment could be sufficient for conferring PHB synthase activity.

This hypothesis was confirmed when we used the screen for PHB production in *A. eutrophus phbC* to identify a smaller complementing region of DNA. In separate reactions, cosmid pLW111-2 was digested with either *Eco*RI or *Bam*HI, religated and transformed into DH5α competent cells. The transformants were mated with

A. eutrophus phbC and transconjugants selected on MM1 fructose containing the appropriate antibiotics. A small percentage of the transconjugants carrying the cosmids which had been treated with *Eco*RI and religated exhibited the opaque white colony morphology associated with PHB production. No opaque white colonies were seen in the transconjugants carrying the cosmids which had been digested with *Bam*HI and religated. We picked 4 PHB-producing strains and mated the cosmids they were carrying into *E. coli* C2110. Restriction analysis revealed that each of the cosmids that was able to direct PHB production contained a 3.7 kb *Eco*RI fragment ligated into the pLAFR1 backbone. One of these cosmids, designated pPHB6, was chosen for further study.

Sequence of R. meliloti phbC

The *R. meliloti phbC* gene was subjected to DNA sequencing. We constructed subclones of the 3.7 kb region to facilitate sequencing from the *Hin*dIII and *Bam*HI sites. The 3648 nt sequence contained one open reading frame of 611 amino acids flanked by two truncated open reading frames, illustrated in Figure 2-4. The complete nucleotide sequence and deduced translations are shown in Figure 2-5.

The central open reading frame (nt 1336-3171) is highly homologous at both the nucleotide and deduced peptide level to PHB polymerases catalogued in the databases maintained by the National Center for Biotechnology Information, and thus we have named the locus *phbC*. The strongest homology is with the *phbC* gene of *R. meliloti* strain 41 (Tombolini et al., 1995). The *phbC* region sequences reported for these two strains overlap for 2518 nt. Of the 45 nucleotide changes, two thirds (30) fall within the presumed *phbC* coding region. Most (24) of these occur at the third position within a codon, leading to only 5 amino acid differences. Although the sequences around the start site are identical, Rm41 has been

hypothesized to use a downstream start site (nt 1438) for an open reading frame 34 amino acids shorter than the one we have predicted for Rm1021. The sequence data seem to support the choice of the upstream start site. The longer open reading frame maximizes homology with previously cloned PHB polymerases. For example, 11 of the 34 N-terminal amino acids of the Rm1021 PhbC sequence are conserved in the *Rhizobium etli* PHA polymerase. In addition, the start codon at nt 1336 is preceded by a strong ribosome binding site (GGAGGA) at -12 to -7.

The *R. meliloti* Rm1021 PhbC deduced protein has 72% identity and 85% similarity with the PhaC deduced protein of *Rhizobium etli*, and 58% identity and 73% similarity with the PhbC deduced protein of *Methylobacterium extorquens*. *R. meliloti* PhbC includes the motif shown for the *A. eutrophus* enzyme to be the site of phosphopantetheine modification (Gerngross et al., 1994). An alignment of *R. meliloti* Rm1021 PhbC with other PHA polymerase proteins is shown in Figure 2-6.

Upstream of the *phbC* open reading frame is a copy of the *Rhizobium*-specific intergenic mosaic element RIME1. The presence of RIME1 upstream of *R. meliloti* strain 41 *phbC* was noted when this 110 nt element was first reported (Østerås et al., 1995). RIME1 contains two large inverted repeats and is discussed in more detail in Chapter 4.

As mentioned above, the sequence reported here contains two truncated open reading frames. In contrast to *A. eutrophus*, where the genes encoding the β-ketothiolase (*phbA*) and acetoacetyl-CoA reductase (*phbB*) activities are adjacent to the gene encoding PHB polymerase, forming the operon *phbCAB*, the open reading frames adjacent to *R. meliloti phbC* are not predicted to be involved in PHB biosynthesis. The divergently transcribed locus 5' to *phbC* is truncated by the *Eco*RI site at 237 amino acids. The translation of this partial ORF is homologous to

proteins which, based on homology with biochemically characterized proteins and the presence of a conserved binding site for the cofactor pyridoxal phosphate, have been classified as aspartate aminotransferases (Sung et al., 1991). Although the *R. meliloti* 5' locus is truncated prior to the expected location of the binding site, the N-terminal homology is strong enough to classify this deduced protein as a member of Class I of pyridoxal-phosphate-dependent aminotransferases. The strongest homology is observed to the deduced peptide of *E coli* ORF f412, which is 51% identical and 69% similar.

Aspartate aminotransferase activity has been studied in *R. meliloti*, and the major isozyme, encoded by *aatA*, is required for symbiotic nitrogen fixation (Rastogi and Watson, 1991). A second isozyme, which is not required for nitrogen fixation, is encoded by *aatB* (Alfano and Kahn, 1993). The deduced peptide of the locus 5' to *R. meliloti phbC* exhibits some homology with AatA (28% identity/50% similarity) and AatB (27% identity/46% similarity), but much less than that observed with *E. coli* F412. The locus 5' to *phbC* may represent another aspartate aminotransferase, or may have different substrate specificity. Other proteins known to belong to this class include tyrosine aminotransferase and aromatic aminotransferase (Sung et al., 1991). To reflect the homology with aspartate aminotransferases while differentiating it from *aatA* and *aatB*, we have provisionally named this locus *aatC*. An alignment of AatC with *R. meliloti* AatA and *E. coli* F412 is presented in Figure 2-7.

Downstream of and transcribed in the same direction as *phbC* is a second truncated open reading frame whose 148 aa deduced peptide is 46% identical and 64% similar to the N-terminus of *E. coli* Sfs1 (Figure 2-8). The C-terminal domain of Sfs1 contains a putative helix-turn-helix DNA binding motif at amino acids 201-220, and Sfs1 is thought to be a regulatory factor involved in sugar fermentation

stimulation (Kawamukai et al., 1991). The *R. meliloti* sequence reported here ends at amino acid 148, upstream of the location of the DNA binding motif in the homologous protein. We have provisionally named this locus *sfs1*.

Construction of insertions in the R. meliloti phbC *gene*

The 3.7 kb *Eco*RI fragment containing the *R. meliloti phbC* gene was subcloned into the broad host range vector pSW213, to produce pLW119. We constructed a restriction map of the 3.7 kb insert (Figure 2-4A) and chose to make mutations at the two *Bam*HI sites and the unique *Hin*dIII site of pLW119 using the kanamycin/neomycin resistance cassette of miniTn*5*Km. The first construct, pLW150, carries an internal deletion/substitution mutation at the two *Bam*HI sites. The second construct, pLW152, carries an insertion mutation at the *Hin*dIII site. These plasmids (Figure 2-4B) were used to disrupt the *R. meliloti* genome and the resulting *phbC* strains were designated Rm9600 and Rm9601.

R. meliloti *strains lacking* phbC *function do not produce PHB*

PHB production was assayed in *R. meliloti* strains, using *A. eutrophus* wild type and *phbC* as controls. As reported above, no PHB production was detected in Rm1021 cultures grown to saturation in rich LB medium (see Figure 2-2A). PHB production was detectable in Rm1021 cultures grown in MM1 fructose medium, and these conditions were adopted for assay of PHB production in *R. meliloti*. No PHB production was detected in *R. meliloti* strains Rm9600 and Rm9601 (Figure 2-2B), indicating that the *phbC* locus disrupted in these mutants is required for the synthesis of PHB by Rm1021.

R. meliloti phbC *mutant strains are Fix*$^+$

To address the question of whether PHB production is required for effective symbiosis between *M. sativa* and derivatives of *R. meliloti* Rm1021, we performed plant inoculation experiments. Alfalfa seedlings were inoculated with *R. meliloti phbC* strains Rm9600, Rm9601 and the appropriate controls. Nodules formed on plants inoculated with Rm9600 or Rm9601 developed with the same timing and appearance as those formed on plants inoculated with wild type bacteria. Our finding that *phbC* strains derived from *R. meliloti* strain Rm1021 are able to induce nitrogen-fixing nodules on alfalfa is consistent with the finding that PHB non-producing mutants of Rm41 are symbiotically effective (Povolo et al., 1994). These two strains utilize different polysaccharides to facilitate invasion of alfalfa, but neither requires PHB synthase activity for biological nitrogen fixation.

The R. meliloti phbC *locus maps to the first symbiotic megaplasmid*

R. meliloti contains three large replicons: the chromosome and two symbiotic megaplasmids (Sobral et al., 1991). The *phbC* locus was mapped to one of these replicons by the method of Finan et al. (Finan et al., 1985). Genomic DNA was prepared from Rm1021, Rm9601 and *Agrobacterium tumefaciens* strains cured of their Ti plasmids and carrying one or neither of the *R. meliloti* megaplasmids. The genomic DNA was digested with *Eco*RI, subjected to electrophoresis and transferred to a Gene Screen Plus membrane. The DNA was probed with the 1.2 kb *Bam*HI fragment of pLW113 containing most of the *R. meliloti phbC* gene. The probe hybridized to a band of 3.7 kb in the lanes carrying DNA from Rm1021 and the *A. tumefaciens* strain harboring pRmeSU47a, and hybridized to a band approximately 2 kb larger in the lane with DNA from *phbC* mutant Rm9601. These results indicate that the *phbC* locus maps to pRmeSU47a and confirm the

construction of the insertion mutation. They also show that *A. tumefaciens* does not contain a chromosomally encoded *phbC* homologue that hybridizes to *R. meliloti phbC* under the stringency conditions used.

To determine whether *phbC* maps near the *nod-nif* region of pRmeSU47a, we used the deletion mapping strains constructed by Truchet et al. (Truchet et al., 1985). These strains contain overlapping deletions covering more than 360 kb surrounding the *nod-nif* region. If an insertion maps to a region which is deleted in the mapping strain, one will be unable to transduce a drug resistance marker from the locus of interest to the mapping strain. The kanamycin resistance marker from Rm9601 was transducible into the deletion mapping strains and their isogenic parent at a similar frequency, indicating that the *phbC* locus does not map to the region defined by these deletions.

DISCUSSION

In this work, we report the isolation of *phbC*, the gene encoding PHB synthase, from *R. meliloti* strain Rm1021. The *phbC* gene has also been characterized in *R. meliloti* isolate Rm41. These two strains are distinctly different, and we wanted to learn whether Rm1021, which has been studied in greater detail than Rm41, encoded a PHB synthase which resembles that of Rm41. Although Rm41 and Rm1021 use dissimilar surface polysaccharides to permit the invasion of indeterminate alfalfa nodules, we found that the *phbC* genes of the two strains are very similar.

Two independent strategies were used to clone the PHB synthase genes from Rm1021 and Rm41. In our work, a cosmid carrying the *phbC* gene of Rm1021 was isolated by heterologous complementation of an *Alcaligenes eutrophus phbC* mutation. After identification of a smaller complementing fragment, defined disruptions were constructed and recombined into the *R. meliloti* genome. These insertions eliminate the ability of the resultant strains to produce PHB, but do not affect the ability of the strain to form nitrogen fixing nodules in association with alfalfa.

Instead of heterologous complementation, a direct screen was employed to identify mutants of Rm41 unable to produce PHB. Povolo *et al.* performed random Tn*5* mutagenesis of *R. meliloti* strain 41 (Povolo et al., 1994). Their initial strategy was to replica plate mutagenized cells and screen with Sudan Black B, a lipophilic dye which stains PHB (Burdon, 1946; Schlegel and Steinbuchel, 1991). They found this method to be inefficient, perhaps because exopolysaccharides produced by wild type *R. meliloti* exclude Sudan Black B from the cell (L. B. Willis and G. C. Walker, unpublished results). Therefore, they screened approximately 1000 mutants by gas chromatography and identified four strains which did not produce PHB. They demonstrated by electron microscopy that the mutants did not contain PHB

granules. The PHB$^-$ mutants were able to induce effective nodules on *Medicago sativa*.

Taken together, the Rm1021 and Rm41 *phbC* results show decisively that PHB production is not required for effective nodulation of *Medicago sativa* by *R. meliloti*. Both of these examples involve indeterminate nodules, in which PHB granules disappear during the normal course of bacteroid development.

Bacteroids in determinate nodules, such as those formed in the interaction between *Rhizobium etli* and bean, accumulate PHB during symbiosis, and a PHB synthase defect leads to a more dramatic effect in this interaction. Cevallos et al. reported the isolation of a *phaC* mutant of *Rhizobium etli*, a symbiont of *Phaseolus vulgaris* (bean) (Cevallos et al., 1996). The PHA synthase mutant grew poorly on pyruvate or glucose as carbon sources and excreted metabolites, including β-hydroxybutyrate, into the growth medium. They showed that, in contrast to the wild type strain, the *phaC* mutants do not accumulate PHB during symbiosis with bean. Far from abolishing nitrogen fixation, this mutation led to increased activity of nitrogenase and a delay in nodule senescence. This result seems to confirm early models that the efficiency of nitrogenase decreased when high levels of PHB are present (Romanov et al., 1974; Tombolini and Nuti, 1989). Although the mutation offers an advantage during symbiosis, the pleiotropic effects may reduce competitiveness of the free-living bacteria in the rhizosphere.

The result that an *R. etli phbC* strain exhibits higher levels of nitrogenase and an extended period of nitrogen fixation (Cevallos et al., 1996) suggests that increasing PHB production in *R. meliloti* could be detrimental to the symbiotic process. The authors reported (Cevallos et al., 1996) that such an experiment was in progress, but no formal report has been made of the results.

The *phbC* gene of *R. meliloti* strain Rm1021 is located on pRmeSU47a, the smaller of two megaplasmids found in this strain. It is worth noting that the open reading frames adjacent to *phbC* are neither *phbA* nor *phbB*. In *Alcaligenes eutrophus*, the genes encoding biosynthetic β-ketothiolase and the NADPH-dependent acetoacyl-CoA reductase are located directly downstream of the PHB synthase gene, forming the operon *phbCAB* (Schubert et al., 1988; Peoples and Sinskey, 1989b; Peoples and Sinskey, 1989c). This pattern is often, but not always, found in other strains of bacteria. For example, *Pseudomonas olerovorans* (Huisman et al., 1991) and *P. aueriginosa* (Timm and Steinbüchel, 1992) each carry two PHA synthase genes in an operon with PHA depolymerase genes. In *Zooglea ramigera*, the PHB synthase gene does not map near the thiolase and reductase genes (Peoples et al., 1987; Peoples and Sinskey, 1989a). The PHB synthase gene of *Thiocystis violacea* is transcribed divergently from the thiolase gene. Although Tombolini et al. have reported the sequence of the *phaA* and *phaB* genes of *R. meliloti* strain Rm41 (Tombolini et al., 1995), no information was submitted about where these loci map with respect to *phbC* or the three *R. meliloti* replicons.

It is interesting that the genetic organization at the *R. meliloti phbC* locus, *aatC phbC sfs1*, is partially conserved at the *phbC* locus in *Methylobacterium extorquens*. The ORF2 locus upstream of *M. extorquens phbC* is homologous to *R. meliloti aatC*, and has also been classified as a putative aspartate aminotransferase. The locus downstream of *M. extorquens phbC* (ORF1) is not closely related to *sfs1*, but encodes a peptide with homology to RosB, a putative regulator of O-antigen expression from *Yersinia enterocolitica* (Zhang et al., 1997) and may therefore be involved in regulation. RosB, ORF1 and Sfs1 all contain putative DNA binding domains. It would be interesting to learn whether the *R. meliloti* locus *sfs1* is cotranscribed with

phbC and to learn both what controls its expression and whether it is a regulatory locus.

PHB production in *R. meliloti* appears to be regulated in the free-living state. In the work reported here, no PHB production was detected when wild type *R. meliloti* was grown to saturation in LB. In order to detect production of PHB by *R. meliloti* strains, we used the nitrogen limiting medium MM1 (Peoples and Sinskey, 1989b). Encarnacion et al. (Encarnacion et al., 1993) found that after *R. meliloti* was subcultured in minimal medium, it underwent unbalanced growth. This phenomenon is characterized by excretion of organic and amino acids into the medium, accumulation of PHB and, after several subcultures, reduction in and cessation of growth. Addition of D-biotin and thiamine to the medium alleviated the unbalanced growth.

A *phbC*::*lacZ* fusion could be employed to identify genes involved in the regulation of *phbC* expression in the free living state. After establishing conditions under which *lacZ* expression is seen in free living cells, a *phbC*::*lacZ* strain would be subjected to chemical or transposon mutagenesis. Transposon mutagenesis could involve transduction of random Tn*5* insertions into a *phbC*::*lacZ* strain, replica plating and screening for altered levels of β-galactosidase activity.

The precise timing of the disappearance of PHB granules during bacteroid development, and the fact that the granules remain inside senescent *bacA* symbiosomes, suggests the disappearance of PHB granules is also regulated *in planta*. It is possible that the *phbC* gene is downregulated after release from the infection thread, that expression or activity of an intracellular PHB depolymerase is increased, or that the substrate concentrations inside the symbiosome favor PHB degradation and not accumulation. Analysis of the expression of a *phbC*::*lacZ* fusion *in planta* could shed light on these questions.

Table 2-1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
Strain		
Rhizobium meliloti		
Rm1021	SU47 *str*-21 (Sm^R)	F. Ausubel
RCR2011	RCR2011 ≡ SU47	(Truchet et al., 1985)
GMI255	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI766	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI956	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI963	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
Rm9600	Rm1021 *phbC*150	this work
Rm9601	Rm1021 *phbC*152	this work
Alcaligenes eutrophus		
H16	wild type, Sm^S	O. Peoples
PHB#2	H16 *phbC*::Tn*5*	O. Peoples
Agrobacterium tumefaciens		
At123	GMI9023 ≡ GMI9050 cured of pAtC58. Sm^R, $Rifampicin^R$	T. M. Finan
At125	GMI9023 pRmeSU47bΩ5007::Tn*5-oriT*	T. M. Finan
At128	GMI9023 pRmeSU47aΩ30::Tn*5-11*	T. M. Finan
Escherichia coli		
C2110	*polA*, Nal^R	B. Staskawicz
DH5α	standard cloning strain	Clontech
Plasmid		
pBluescript SK+	phagemid vector, Amp^R	Stratagene
pRK600	pRK2013 *npt*::Tn9 Cm^R	(Finan et al., 1984)
pLAFR1	Tc^R, Mobilizable RK2 cosmid	(Friedman et al., 1982)
pPH1JI	Gm^R, Sp^R, IncP	(Beringer et al., 1978)
pSW213	Tc^R, IncP broad host range vector	(Chen and Winans, 1991)
pUTminiTn*5*Km	source of Km^R/Nm^R cassette	K. Timmis (de Lorenzo et al., 1990)
pLW111-1	Tc^R derivative of pLAFR1 containing *phbC* locus of *R. meliloti*	this work
pLW111-2	Tc^R derivative of pLAFR1 containing *phbC* locus of *R. meliloti*	this work
pPHB6	pLAFR1 containing 3.8 kb *Eco*RI fragment from pLW111	this work
pLW119	pSW213 containing 3.8 kb *Eco*RI fragment from pPHB6	this work
pLW150	pLW119 containing Km^R/Nm^R cassette in place of 1.2 kb *Bam*HI fragment	this work
pLW152	pLW119 containing Km^R/Nm^R cassette at *Hin*dIII site	this work

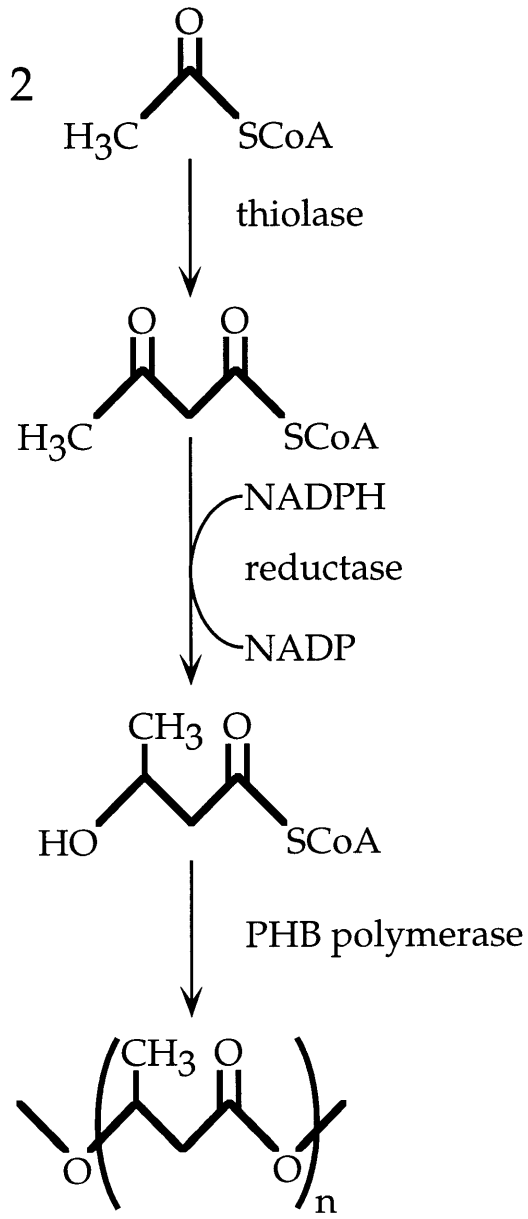


Figure 2-1.
PHB biosynthetic pathway in *Alcaligenes eutrophus*. Adapted from Peoples and Sinskey (Peoples and Sinskey, 1989b).

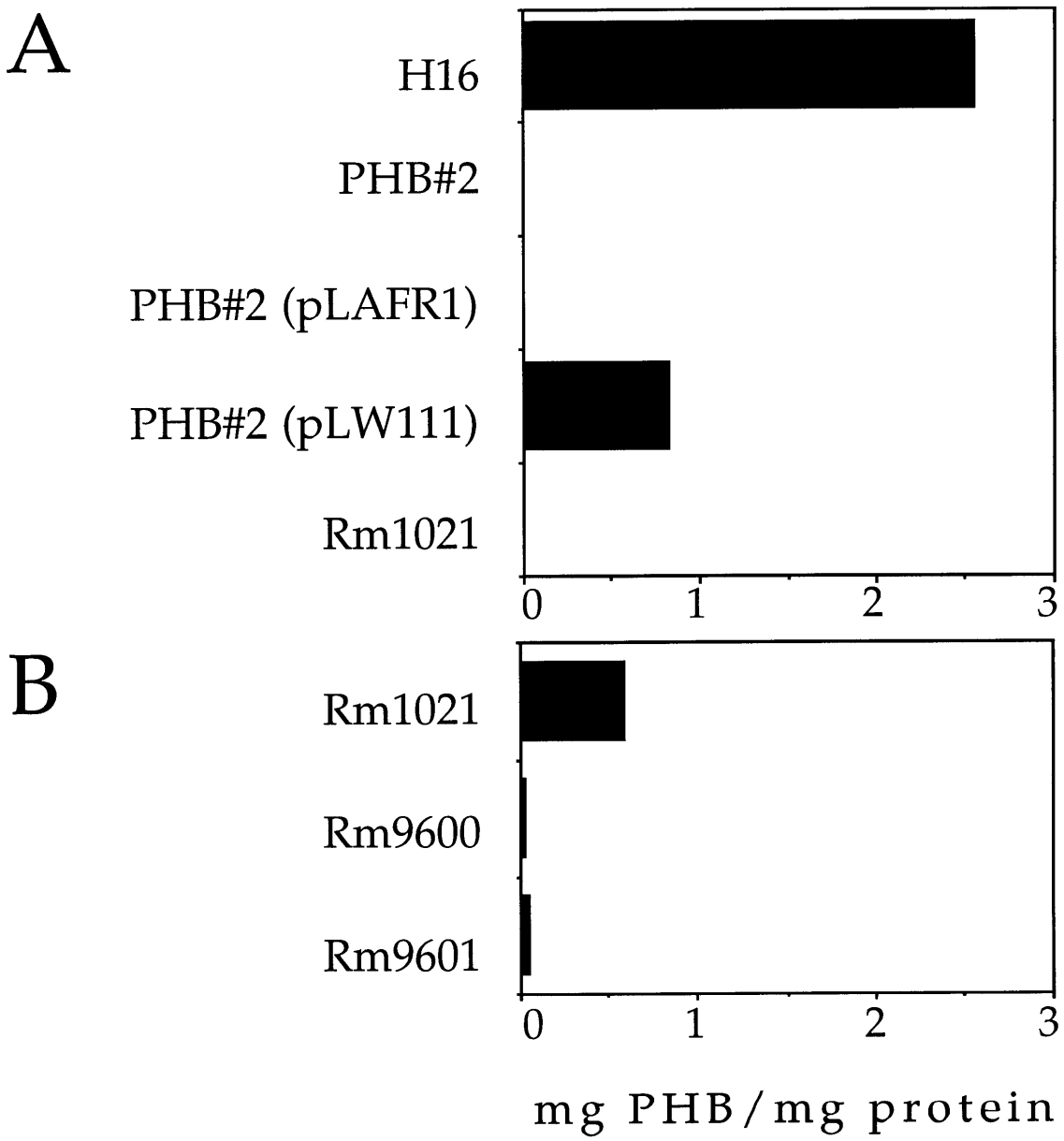


Figure 2-2.
(A) PHB production by *A. eutrophus* strains and wild type Rm1021 grown in LB medium. Lane 1: wild type *A. eutrophus*. Lanes 2-4: *A. eutrophus phbC* harboring no plasmid (lane 2), pLAFR1 (lane 3) or pLW111-1 (lane 4). Lane 5: *R. meliloti* Rm1021.

(B) PHB production by *R. meliloti* strains grown in MM1 fructose medium. Lane 1: Rm1021. Lane 2: Rm9600. Lane 3: Rm9601.

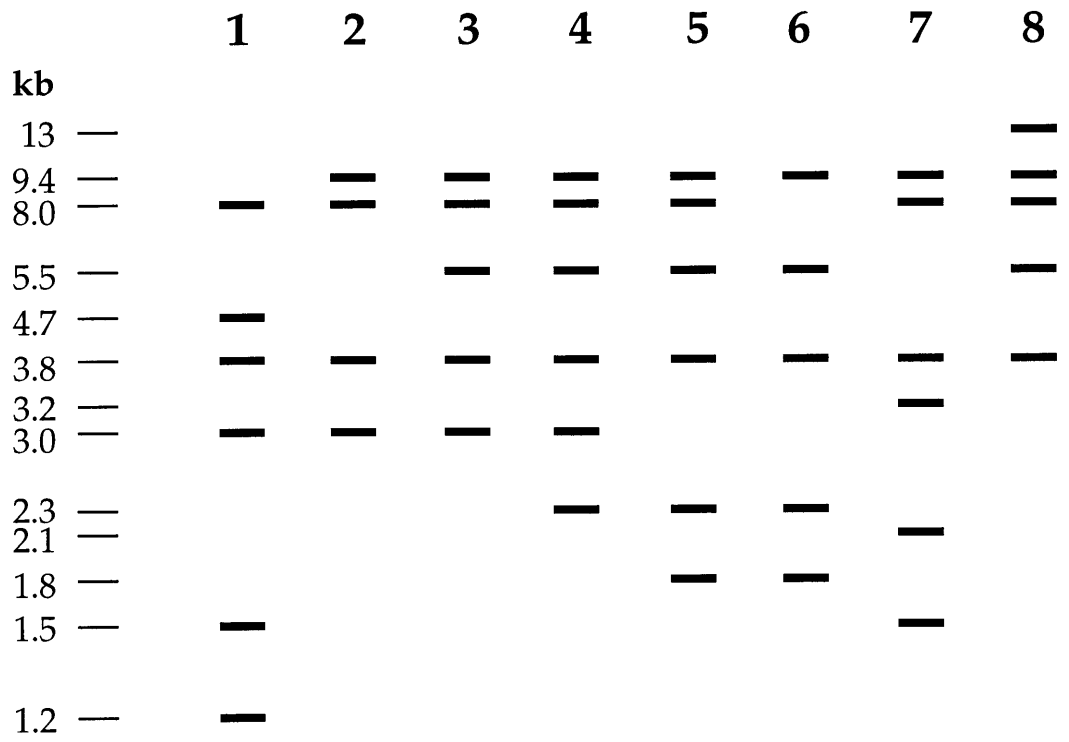


Figure 2-3.
Illustration depicting the eight different restriction patterns of inserts from PHB$^+$ cosmids identified in this study. The illustration was constructed using data gathered from electrophoretic separation of *Eco*RI-digested cosmid DNA in 0.6% agarose.

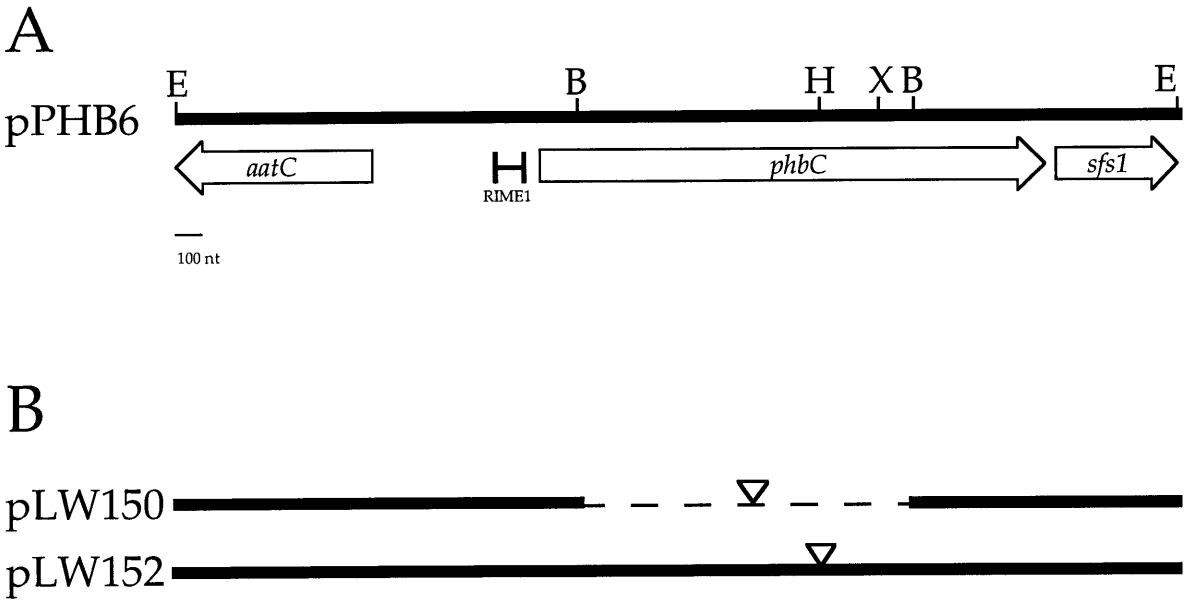
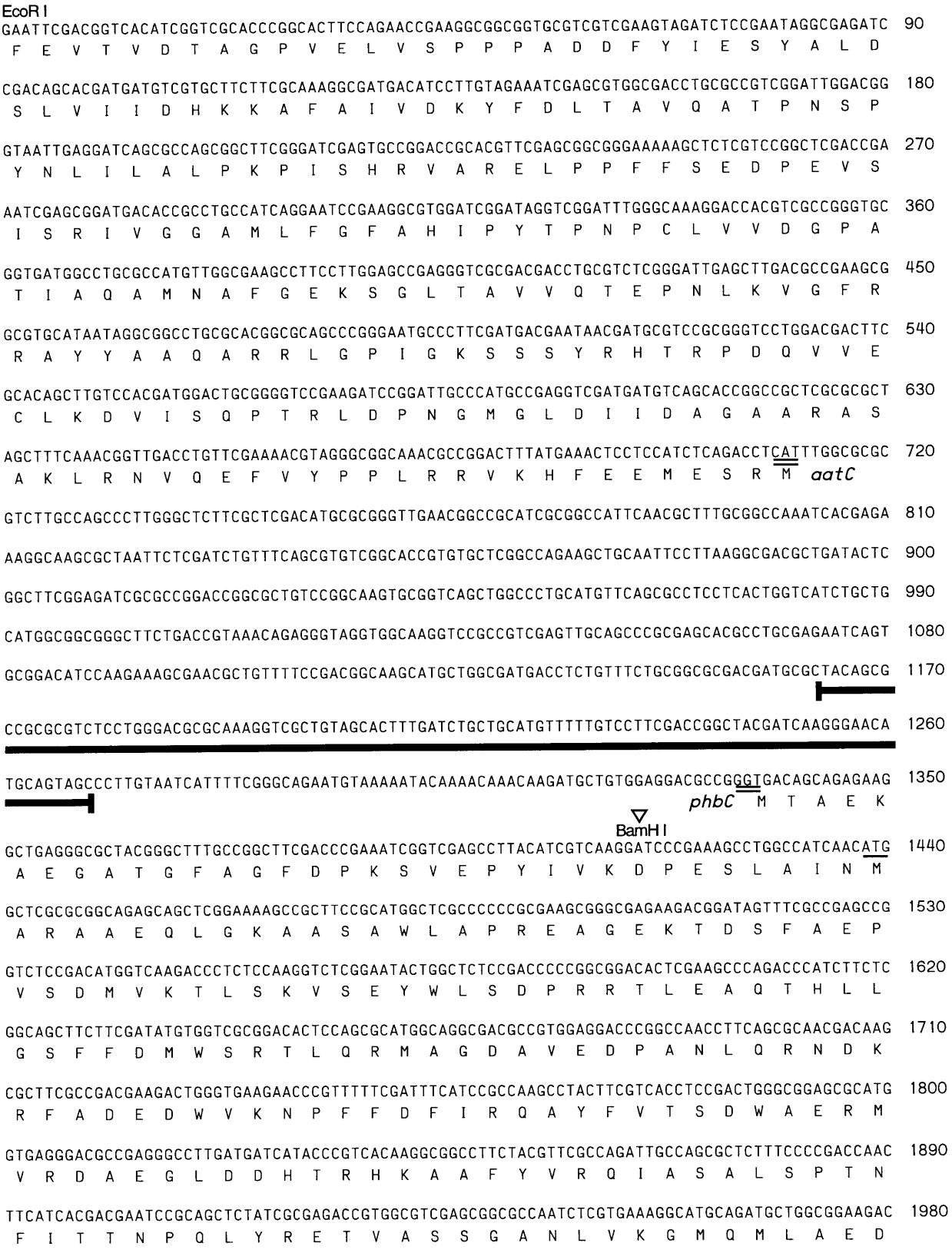


Figure 2-4.
Genetic organization of the *R. meliloti* Rm1021 *phbC* region. (A) Restriction map of the insert of pPHB6. The location and direction of transcription of *aatC*, *phbC* and *sfs1* are indicated by open arrows. The position of RIME1 is shown. Restriction sites shown are: *Eco*RI (E), *Bam*HI (B), *Hin*dIII (H), *Xho*I (X).

(B) Derivatives of pPHB6 with disruptions in the *R. meliloti phbC* gene. Scale and restriction sites are the same as in panel A. Downward triangles indicate insertion of the kanamycin resistance marker from miniTn*5*Km. The *R. meliloti* DNA which is deleted in pLW150 is indicated with a dashed line.

Figure 2-5.
Nucleotide sequence of the phbC gene of *R. meliloti* strain Rm1021. The sequence and deduced polypeptides for *phbC* and the truncated open reading frames *aatC* and *sfs1* are shown. The presumed initiation codons are double underscored. An alternative start codon for *phbC* is underscored. The sites of kanamycin resistance cassette insertions are marked with downward arrows.


EcoR I
GAATTCGACGGTCACATCGGTCGCACCCGGCACTTCCAGAACCGAAGGCGGCGGTGCGTCGTCGAAGTAGATCTCCGAATAGGCGAGATC 90
F E V T V D T A G P V E L V S P P P A D D F Y I E S Y A L D
CGACAGCACGATGATGTCGTGCTTCTTCGCAAAGGCGATGACATCCTTGTAGAAATCGAGCGTGGCGACCTGCGCCGTCGGATTGGACGG 180
S L V I I D H K K A F A I V D K Y F D L T A V Q A T P N S P
GTAATTGAGGATCAGCGCCAGCGGCTTCGGGATCGAGTGCCGGACCGCACGTTCGAGCGGCGGGAAAAAGCTCTCGTCCGGCTCGACCGA 270
Y N L I L A L P K P I S H R V A R E L P P F F S E D P E V S
AATCGAGCGGATGACACCGCCTGCCATCAGGAATCCGAAGGCGTGGATCGGATAGGTCGGATTTGGGCAAAGGACCACGTCGCCGGGTGC 360
I S R I V G G A M L F G F A H I P Y T P N P C L V V D G P A
GGTGATGGCCTGCGCCATGTTGGCGAAGCCTTCCTTGGAGCCGAGGGTCGCGACGACCTGCGTCTCGGGATTGAGCTTGACGCCGAAGCG 450
T I A Q A M N A F G E K S G L T A V V Q T E P N L K V G F R
GCGTGCATAATAGGCGGCCTGCGCACGGCGCAGCCCGGGAATGCCCTTCGATGACGAATAACGATGCGTCCGCGGGTCCTGGACGACTTC 540
R A Y Y A A Q A R R L G P I G K S S S Y R H T R P D Q V V E
GCACAGCTTGTCCACGATGGACTGCGGGGTCCGAAGATCCGGATTGCCCATGCCGAGGTCGATGATGTCAGCACCGGCCGCTCGCGCGCT 630
C L K D V I S Q P T R L D P N G M G L D I I D A G A A R A S
AGCTTTCAAACGGTTGACCTGTTCGAAAACGTAGGGCGGCAAACGCCGGACTTTATGAAACTCCTCCATCTCAGACCTCATTTGGCGCGC 720
A K L R N V Q E F V Y P P L R R V K H F E E M E S R M aatC
GTCTTGCCAGCCCTTGGGCTCTTCGCTCGACATGCGCGGGTTGAACGGCCGCATCGCGGCCATTCAACGCTTTGCGGCCAAATCACGAGA 810
AAGGCAAGCGCTAATTCTCGATCTGTTTCAGCGTGTCGGCACCGTGTGCTCGGCCAGAAGCTGCAATTCCTTAAGGCGACGCTGATACTC 900
GGCTTCGGAGATCGCGCCGGACCGGCGCTGTCCGGCAAGTGCGGTCAGCTGGCCCTGCATGTTCAGCGCCTCCTCACTGGTCATCTGCTG 990
CATGGCGGCGGGCTTCTGACCGTAAACAGAGGGTAGGTGGCAAGGTCCGCCGTCGAGTTGCAGCCCGCGAGCACGCCTGCGAGAATCAGT 1080
GCGGACATCCAAGAAAGCGAACGCTGTTTTCCGACGGCAAGCATGCTGGCGATGACCTCTGTTTCTGCGGCGCGACGATGCGCTACAGCG 1170
CCGCGCGTCTCCTGGGACGCGCAAAGGTCGCTGTAGCACTTTGATCTGCTGCATGTTTTTGTCCTTCGACCGGCTACGATCAAGGGAACA 1260
TGCAGTAGCCCTTGTAATCATTTTCGGGCAGAATGTAAAAATACAAAACAAACAAGATGCTGTGGAGGACGCCGGGTGACAGCAGAGAAG 1350
phbC M T A E K
BamH I
GCTGAGGGCGCTACGGGCTTTGCCGGCTTCGACCCGAAATCGGTCGAGCCTTACATCGTCAAGGATCCCGAAAGCCTGGCCATCAACATG 1440
A E G A T G F A G F D P K S V E P Y I V K D P E S L A I N M
GCTCGCGCGGCAGAGCAGCTCGGAAAAGCCGCTTCCGCATGGCTCGCCCCCCGCGAAGCGGGCGAGAAGACGGATAGTTTCGCCGAGCCG 1530
A R A A E Q L G K A A S A W L A P R E A G E K T D S F A E P
GTCTCCGACATGGTCAAGACCCTCTCCAAGGTCTCGGAATACTGGCTCTCCGACCCCCGGCGGACACTCGAAGCCCAGACCCATCTTCTC 1620
V S D M V K T L S K V S E Y W L S D P R R T L E A Q T H L L
GGCAGCTTCTTCGATATGTGGTCGCGGACACTCCAGCGCATGGCAGGCGACGCCGTGGAGGACCCGGCCAACCTTCAGCGCAACGACAAG 1710
G S F F D M W S R T L Q R M A G D A V E D P A N L Q R N D K
CGCTTCGCCGACGAAGACTGGGTGAAGAACCCGTTTTTCGATTTCATCCGCCAAGCCTACTTCGTCACCTCCGACTGGGCGGAGCGCATG 1800
R F A D E D W V K N P F F D F I R Q A Y F V T S D W A E R M
GTGAGGGACGCCGAGGGCCTTGATGATCATACCCGTCACAAGGCGGCCTTCTACGTTCGCCAGATTGCCAGCGCTCTTTCCCCGACCAAC 1890
V R D A E G L D D H T R H K A A F Y V R Q I A S A L S P T N
TTCATCACGACGAATCCGCAGCTCTATCGCGAGACCGTGGCGTCGAGCGGCGCCAATCTCGTGAAAGGCATGCAGATGCTGGCGGAAGAC 1980
F I T T N P Q L Y R E T V A S S G A N L V K G M Q M L A E D

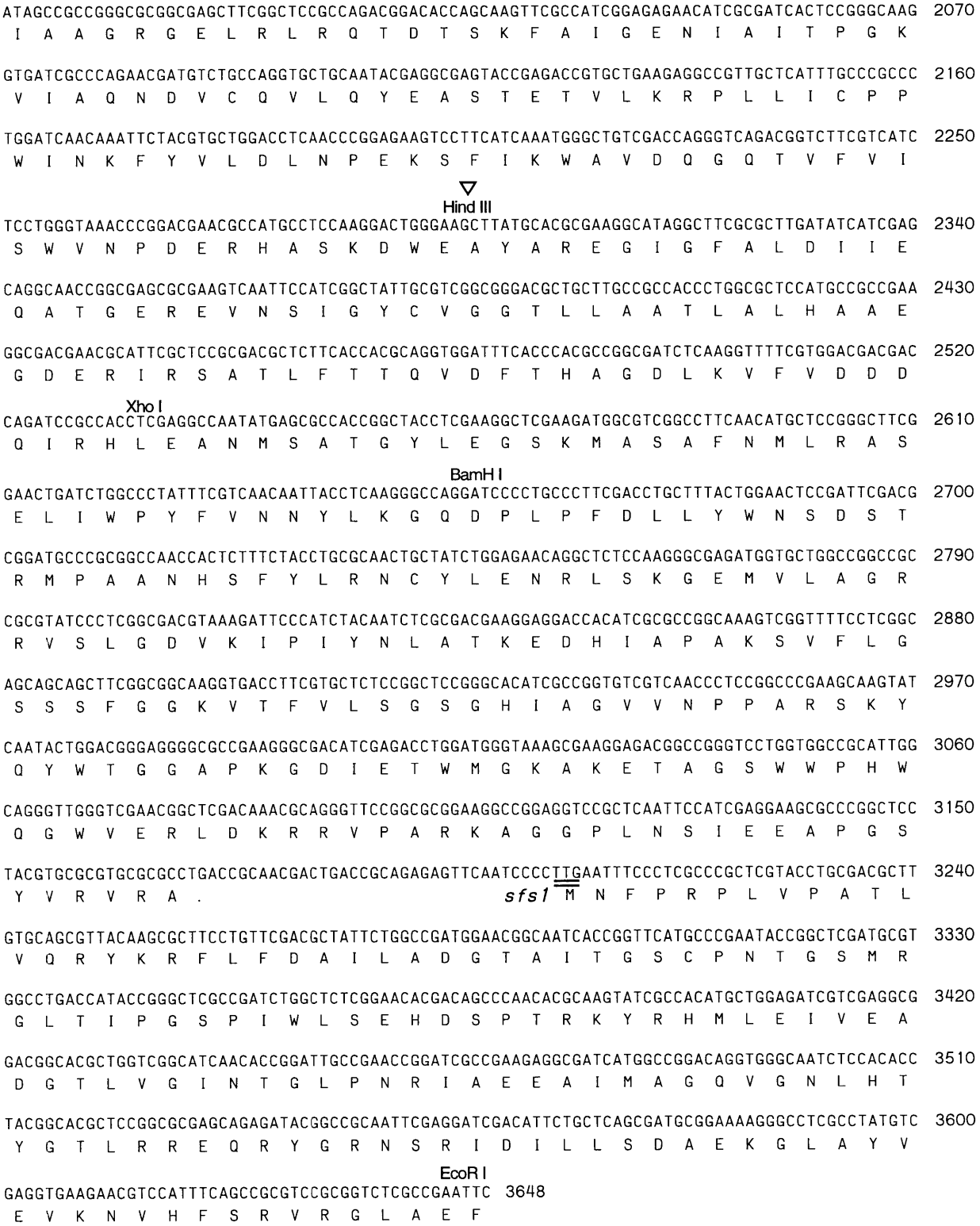
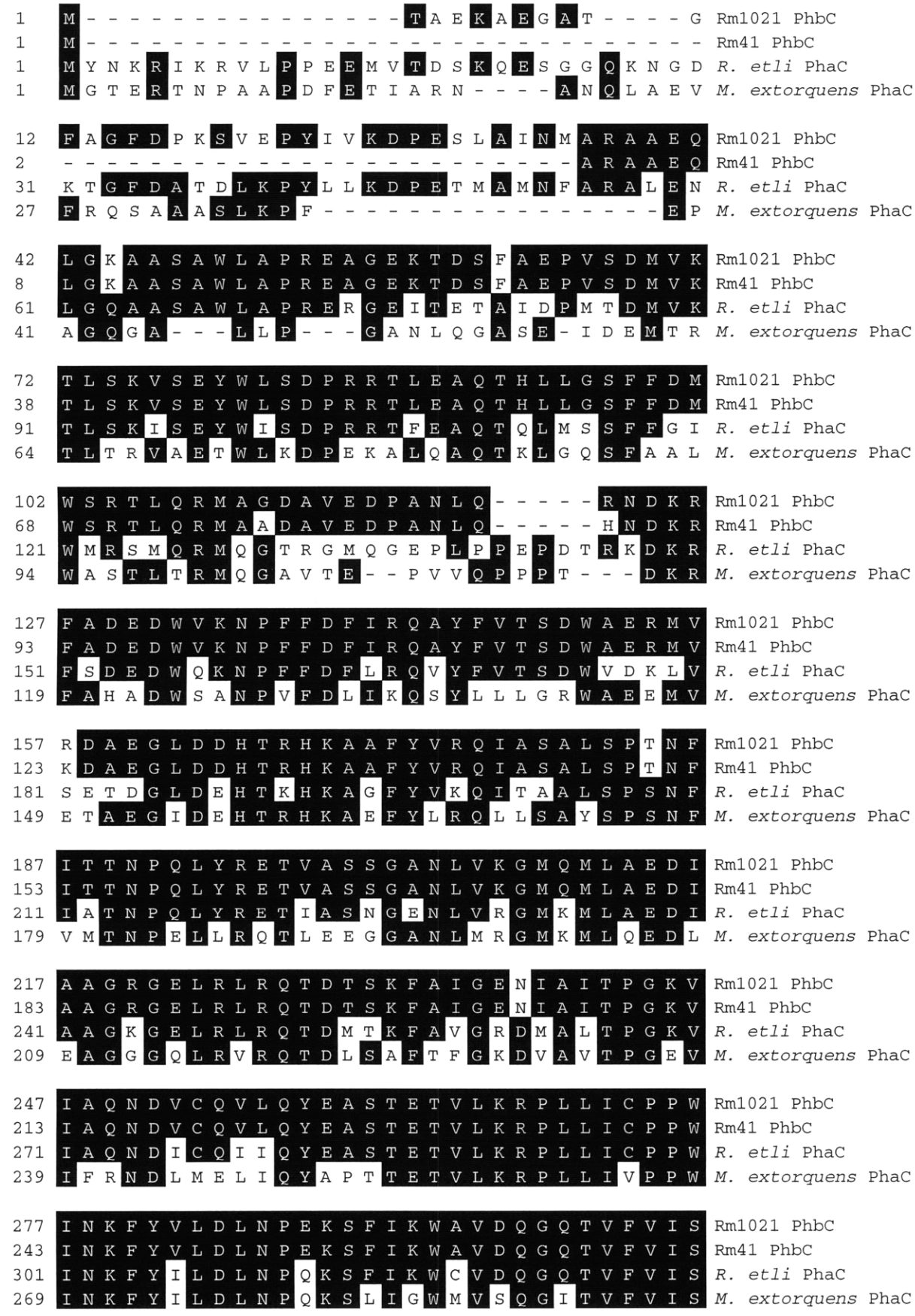
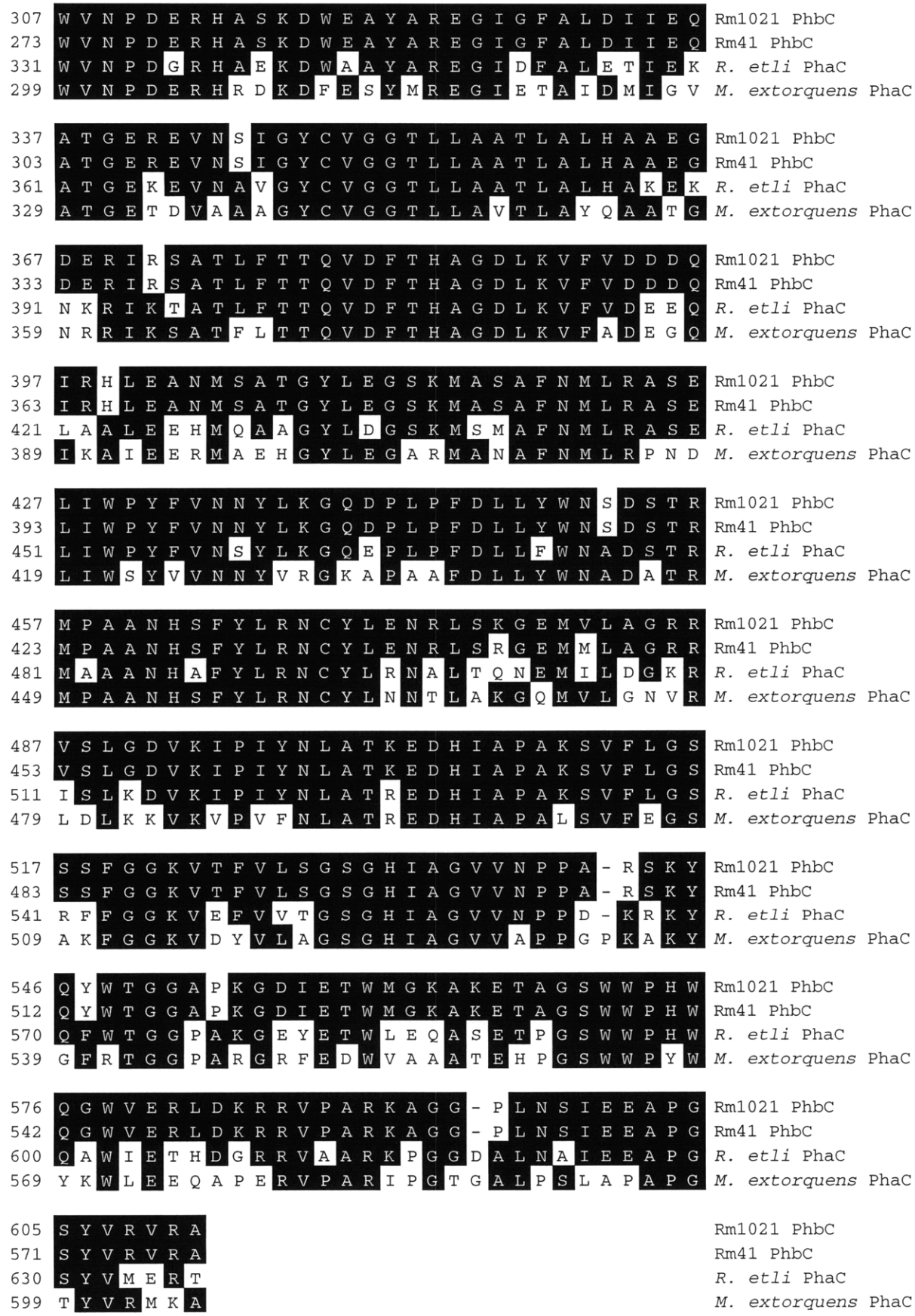

ATAGCCGCCGGGCGCGGCGAGCTTCGGCTCCGCCAGACGGACACCAGCAAGTTCGCCATCGGAGAGAACATCGCGATCACTCCGGGCAAG 2070
I A A G R G E L R L R Q T D T S K F A I G E N I A I T P G K
GTGATCGCCCAGAACGATGTCTGCCAGGTGCTGCAATACGAGGCGAGTACCGAGACCGTGCTGAAGAGGCCGTTGCTCATTTGCCCGCCC 2160
V I A Q N D V C Q V L Q Y E A S T E T V L K R P L L I C P P
TGGATCAACAAATTCTACGTGCTGGACCTCAACCCGGAGAAGTCCTTCATCAAATGGGCTGTCGACCAGGGTCAGACGGTCTTCGTCATC 2250
W I N K F Y V L D L N P E K S F I K W A V D Q G Q T V F V I
Hind III
TCCTGGGTAAACCCGGACGAACGCCATGCCTCCAAGGACTGGGAAGCTTATGCACGCGAAGGCATAGGCTTCGCGCTTGATATCATCGAG 2340
S W V N P D E R H A S K D W E A Y A R E G I G F A L D I I E
CAGGCAACCGGCGAGCGCGAAGTCAATTCCATCGGCTATTGCGTCGGCGGGACGCTGCTTGCCGCCACCCTGGCGCTCCATGCCGCCGAA 2430
Q A T G E R E V N S I G Y C V G G T L L A A T L A L H A A E
GGCGACGAACGCATTCGCTCCGCGACGCTCTTCACCACGCAGGTGGATTTCACCCACGCCGGCGATCTCAAGGTTTTCGTGGACGACGAC 2520
G D E R I R S A T L F T T Q V D F T H A G D L K V F V D D D
Xho I
CAGATCCGCCACCTCGAGGCCAATATGAGCGCCACCGGCTACCTCGAAGGCTCGAAGATGGCGTCGGCCTTCAACATGCTCCGGGCTTCG 2610
Q I R H L E A N M S A T G Y L E G S K M A S A F N M L R A S
BamH I
GAACTGATCTGGCCCTATTTCGTCAACAATTACCTCAAGGGCCAGGATCCCCTGCCCTTCGACCTGCTTTACTGGAACTCCGATTCGACG 2700
E L I W P Y F V N N Y L K G Q D P L P F D L L Y W N S D S T
CGGATGCCCGCGGCCAACCACTCTTTCTACCTGCGCAACTGCTATCTGGAGAACAGGCTCTCCAAGGGCGAGATGGTGCTGGCCGGCCGC 2790
R M P A A N H S F Y L R N C Y L E N R L S K G E M V L A G R
CGCGTATCCCTCGGCGACGTAAAGATTCCCATCTACAATCTCGCGACGAAGGAGGACCACATCGCGCCGGCAAAGTCGGTTTTCCTCGGC 2880
R V S L G D V K I P I Y N L A T K E D H I A P A K S V F L G
AGCAGCAGCTTCGGCGGCAAGGTGACCTTCGTGCTCTCCGGCTCCGGGCACATCGCCGGTGTCGTCAACCCTCCGGCCCGAAGCAAGTAT 2970
S S S F G G K V T F V L S G S G H I A G V V N P P A R S K Y
CAATACTGGACGGGAGGGGCGCCGAAGGGCGACATCGAGACCTGGATGGGTAAAGCGAAGGAGACGGCCGGGTCCTGGTGGCCGCATTGG 3060
Q Y W T G G A P K G D I E T W M G K A K E T A G S W W P H W
CAGGGTTGGGTCGAACGGCTCGACAAACGCAGGGTTCCGGCGCGGAAGGCCGGAGGTCCGCTCAATTCCATCGAGGAAGCGCCCGGCTCC 3150
Q G W V E R L D K R R V P A R K A G G P L N S I E E A P G S
TACGTGCGCGTGCGCGCCTGACCGCAACGACTGACCGCAGAGAGTTCAATCCCCTTGAATTTCCCTCGCCCGCTCGTACCTGCGACGCTT 3240
Y V R V R A . sfs1 M N F P R P L V P A T L
GTGCAGCGTTACAAGCGCTTCCTGTTCGACGCTATTCTGGCCGATGGAACGGCAATCACCGGTTCATGCCCGAATACCGGCTCGATGCGT 3330
V Q R Y K R F L F D A I L A D G T A I T G S C P N T G S M R
GGCCTGACCATACCGGGCTCGCCGATCTGGCTCTCGGAACACGACAGCCCAACACGCAAGTATCGCCACATGCTGGAGATCGTCGAGGCG 3420
G L T I P G S P I W L S E H D S P T R K Y R H M L E I V E A
GACGGCACGCTGGTCGGCATCAACACCGGATTGCCGAACCGGATCGCCGAAGAGGCGATCATGGCCGGACAGGTGGGCAATCTCCACACC 3510
D G T L V G I N T G L P N R I A E E A I M A G Q V G N L H T
TACGGCACGCTCCGGCGCGAGCAGAGATACGGCCGCAATTCGAGGATCGACATTCTGCTCAGCGATGCGGAAAAGGGCCTCGCCTATGTC 3600
Y G T L R R E Q R Y G R N S R I D I L L S D A E K G L A Y V
EcoR I
GAGGTGAAGAACGTCCATTTCAGCCGCGTCCGCGGTCTCGCCGAATTC 3648
E V K N V H F S R V R G L A E F

Figure 2-6.

Alignment of the deduced peptides of *R. meliloti* strain Rm1021 PHB synthase with the PHB synthase of *R. meliloti* strain Rm41 (g790554), the PHA synthase from *R. etli* (g1209028), and the PHB synthase from *M. extorquens* (g349493). Protein sequence accession numbers are indicated in parentheses. Conserved residues are shaded with black.


1 M - - - - - - - - - - - - - - - - T A E K A E G A T - - - - G Rm1021 PhbC
1 M - Rm41 PhbC
1 M Y N K R I K R V L P P E E M V T D S K Q E S G G Q K N G D R. etli PhaC
1 M G T E R T N P A A P D F E T I A R N - - - - A N Q L A E V M. extorquens PhaC
12 F A G F D P K S V E P Y I V K D P E S L A I N M A R A A E Q Rm1021 PhbC
2 - A R A A E Q Rm41 PhbC
31 K T G F D A T D L K P Y L L K D P E T M A M N F A R A L E N R. etli PhaC
27 F R Q S A A A S L K P F - - - - - - - - - - - - - - - - E P M. extorquens PhaC
42 L G K A A S A W L A P R E A G E K T D S F A E P V S D M V K Rm1021 PhbC
8 L G K A A S A W L A P R E A G E K T D S F A E P V S D M V K Rm41 PhbC
61 L G Q A A S A W L A P R E R G E I T E T A I D P M T D M V K R. etli PhaC
41 A G Q G A - - - L L P - - - G A N L Q G A S E - I D E M T R M. extorquens PhaC
72 T L S K V S E Y W L S D P R R T L E A Q T H L L G S F F D M Rm1021 PhbC
38 T L S K V S E Y W L S D P R R T L E A Q T H L L G S F F D M Rm41 PhbC
91 T L S K I S E Y W I S D P R R T F E A Q T Q L M S S F F G I R. etli PhaC
64 T L T R V A E T W L K D P E K A L Q A Q T K L G Q S F A A L M. extorquens PhaC
102 W S R T L Q R M A G D A V E D P A N L Q - - - - - R N D K R Rm1021 PhbC
68 W S R T L Q R M A A D A V E D P A N L Q - - - - - H N D K R Rm41 PhbC
121 W M R S M Q R M Q G T R G M Q G E P L P P E P D T R K D K R R. etli PhaC
94 W A S T L T R M Q G A V T E - - P V V Q P P P T - - - D K R M. extorquens PhaC
127 F A D E D W V K N P F F D F I R Q A Y F V T S D W A E R M V Rm1021 PhbC
93 F A D E D W V K N P F F D F I R Q A Y F V T S D W A E R M V Rm41 PhbC
151 F S D E D W Q K N P F F D F L R Q V Y F V T S D W V D K L V R. etli PhaC
119 F A H A D W S A N P V F D L I K Q S Y L L L G R W A E E M V M. extorquens PhaC
157 R D A E G L D D H T R H K A A F Y V R Q I A S A L S P T N F Rm1021 PhbC
123 K D A E G L D D H T R H K A A F Y V R Q I A S A L S P T N F Rm41 PhbC
181 S E T D G L D E H T K H K A G F Y V K Q I T A A L S P S N F R. etli PhaC
149 E T A E G I D E H T R H K A E F Y L R Q L L S A Y S P S N F M. extorquens PhaC
187 I T T N P Q L Y R E T V A S S G A N L V K G M Q M L A E D I Rm1021 PhbC
153 I T T N P Q L Y R E T V A S S G A N L V K G M Q M L A E D I Rm41 PhbC
211 I A T N P Q L Y R E T I A S N G E N L V R G M K M L A E D I R. etli PhaC
179 V M T N P E L L R Q T L E E G G A N L M R G M K M L Q E D L M. extorquens PhaC
217 A A G R G E L R L R Q T D T S K F A I G E N I A I T P G K V Rm1021 PhbC
183 A A G R G E L R L R Q T D T S K F A I G E N I A I T P G K V Rm41 PhbC
241 A A G K G E L R L R Q T D M T K F A V G R D M A L T P G K V R. etli PhaC
209 E A G G G Q L R V R Q T D L S A F T F G K D V A V T P G E V M. extorquens PhaC
247 I A Q N D V C Q V L Q Y E A S T E T V L K R P L L I C P P W Rm1021 PhbC
213 I A Q N D V C Q V L Q Y E A S T E T V L K R P L L I C P P W Rm41 PhbC
271 I A Q N D I C Q I I Q Y E A S T E T V L K R P L L I C P P W R. etli PhaC
239 I F R N D L M E L I Q Y A P T T E T V L K R P L L I V P P W M. extorquens PhaC
277 I N K F Y V L D L N P E K S F I K W A V D Q G Q T V F V I S Rm1021 PhbC
243 I N K F Y V L D L N P E K S F I K W A V D Q G Q T V F V I S Rm41 PhbC
301 I N K F Y I L D L N P Q K S F I K W C V D Q G Q T V F V I S R. etli PhaC
269 I N K F Y I L D L N P Q K S L I G W M V S Q G I T V F V I S M. extorquens PhaC


307 WVNPDERHASKDWEAYAREGIGFALDIIEQ Rm1021 PhbC
273 WVNPDERHASKDWEAYAREGIGFALDIIEQ Rm41 PhbC
331 WVNPDGRHAEKDWAAYAREGIDFALETIEK R. etli PhaC
299 WVNPDERHRDKDFESYMREGIETAIDMIGV M. extorquens PhaC
337 ATGEREVNSIGYCVGGTLLAATLALHAAEG Rm1021 PhbC
303 ATGEREVNSIGYCVGGTLLAATLALHAAEG Rm41 PhbC
361 ATGEKEVNAVGYCVGGTLLAATLALHAKEK R. etli PhaC
329 ATGETDVAAAGYCVGGTLLAVTLAYQAATG M. extorquens PhaC
367 DERIRSATLFTTQVDFTHAGDLKVFVDDDQ Rm1021 PhbC
333 DERIRSATLFTTQVDFTHAGDLKVFVDDDQ Rm41 PhbC
391 NKRIKTATLFTTQVDFTHAGDLKVFVDEEQ R. etli PhaC
359 NRRIKSATFLTTQVDFTHAGDLKVFADEGQ M. extorquens PhaC
397 IRHLEANMSATGYLEGSKMASAFNMLRASE Rm1021 PhbC
363 IRHLEANMSATGYLEGSKMASAFNMLRASE Rm41 PhbC
421 LAALEEHMQAAGYLDGSKMSMAFNMLRASE R. etli PhaC
389 IKAIEERMAEHGYLEGARMANAFNMLRPND M. extorquens PhaC
427 LIWPYFVNNYLKGQDPLPFDLLYWNSDSTR Rm1021 PhbC
393 LIWPYFVNNYLKGQDPLPFDLLYWNSDSTR Rm41 PhbC
451 LIWPYFVNSYLKGQEPLPFDLLFWNADSTR R. etli PhaC
419 LIWSYVVNNYVRGKAPAAFDLLYWNADATR M. extorquens PhaC
457 MPAANHSFYLRNCYLENRLSKGEMVLAGRR Rm1021 PhbC
423 MPAANHSFYLRNCYLENRLSRGEMMLAGRR Rm41 PhbC
481 MAAANHAFYLRNCYLRNALTQNEMILDGKR R. etli PhaC
449 MPAANHSFYLRNCYLNNTLAKGQMVLGNVR M. extorquens PhaC
487 VSLGDVKIPIYNLATKEDHIAPAKSVFLGS Rm1021 PhbC
453 VSLGDVKIPIYNLATKEDHIAPAKSVFLGS Rm41 PhbC
511 ISLKDVKIPIYNLATREDHIAPAKSVFLGS R. etli PhaC
479 LDLKKVKVPVFNLATREDHIAPALSVFEGS M. extorquens PhaC
517 SSFGGKVTFVLSGSGHIAGVVNPPA-RSKY Rm1021 PhbC
483 SSFGGKVTFVLSGSGHIAGVVNPPA-RSKY Rm41 PhbC
541 RFFGGKVEFVVTGSGHIAGVVNPPD-KRKY R. etli PhaC
509 AKFGGKVDYVLAGSGHIAGVVAPPGPKAKY M. extorquens PhaC
546 QYWTGGAPKGDIETWMGKAKETAGSWWPHW Rm1021 PhbC
512 QYWTGGAPKGDIETWMGKAKETAGSWWPHW Rm41 PhbC
570 QFWTGGPAKGEYETWLEQASETPGSWWPHW R. etli PhaC
539 GFRTGGPARGRFEDWVAAATEHPGSWWPYW M. extorquens PhaC
576 QGWVERLDKRRVPARKAGG-PLNSIEEAPG Rm1021 PhbC
542 QGWVERLDKRRVPARKAGG-PLNSIEEAPG Rm41 PhbC
600 QAWIETHDGRRVAARKPGGDALNAIEEAPG R. etli PhaC
569 YKWLEEQAPERVPARIPGTGALPSLAPAPG M. extorquens PhaC
605 SYVRVRA Rm1021 PhbC
571 SYVRVRA Rm41 PhbC
630 SYVMERT R. etli PhaC
599 TYVRMKA M. extorquens PhaC

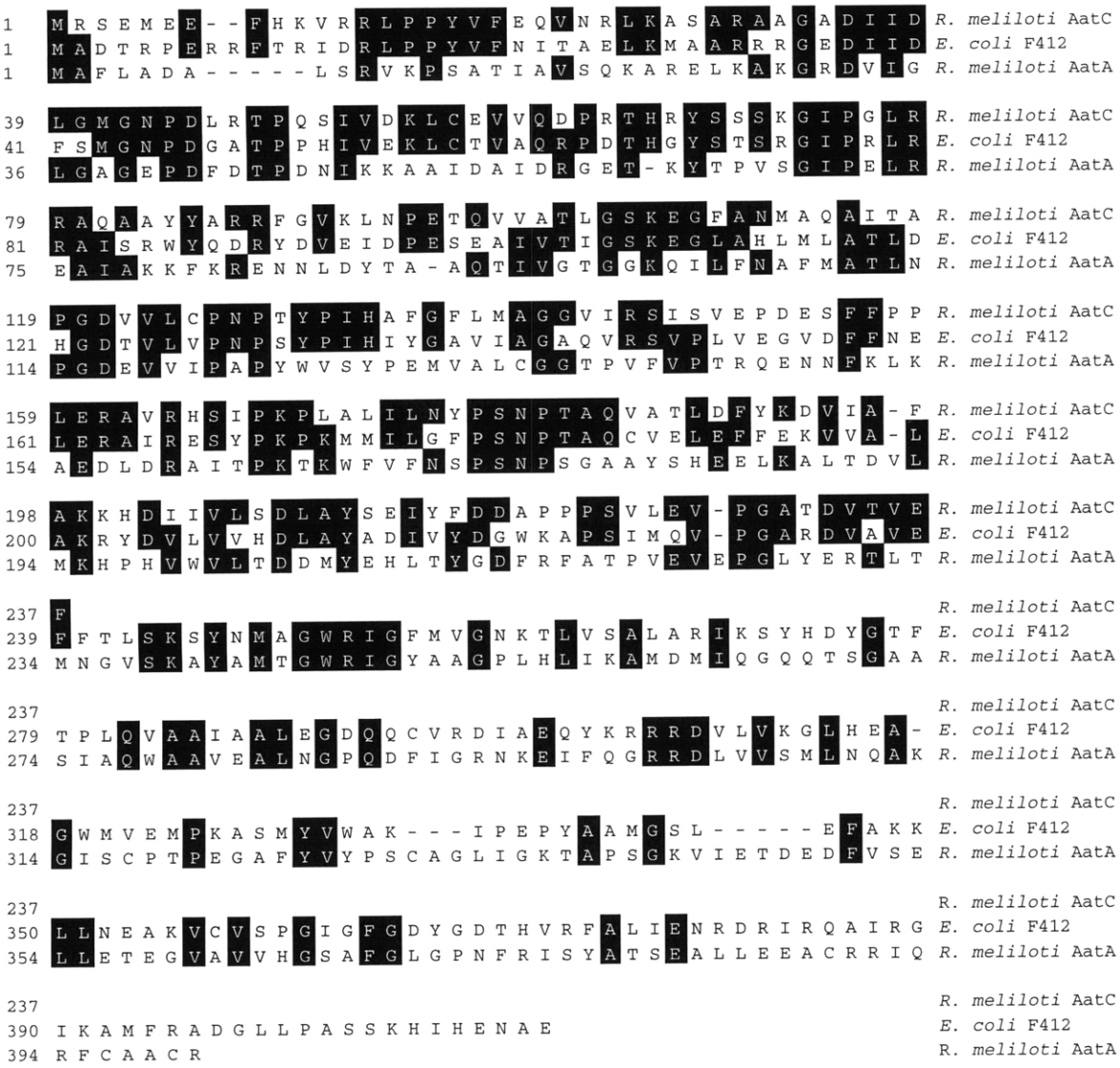


Figure 2-7.

Alignment of presumed aminotransferase deduced peptides *E. coli* ORF f412 (g1788722), *R. meliloti* AatC and *R. meliloti* AatA (g152149). Protein accession numbers are indicated in parentheses. Conserved residues are shaded with black.

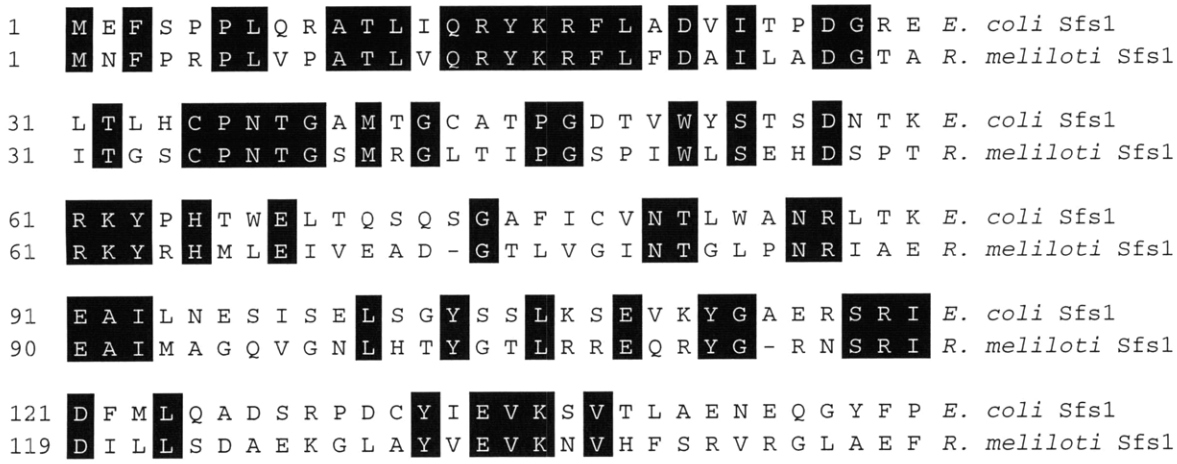


Figure 2-8.

Alignment of the N-termini of deduced peptides *E. coli* Sfs1 (Accession number P18273) and *R. meliloti* Sfs1. Identical residues are shaded with black.

Chapter 3:

Identification of the *Rhizobium meliloti* alcohol dehydrogenase (*adhA*) gene and heterologous expression in *Alcaligenes eutrophus*

ABSTRACT

Sucrose is the primary sugar transported in the phloem and the first carbon-containing compound to be transported into the root nodules and symbiosome organelles of nodulated leguminous plants. In order to identify *Rhizobium meliloti* genes involved in sucrose uptake or hydrolysis, we mated a *R. meliloti* genomic library into an *Alcaligenes eutrophus* strain defective in production of poly-β-hydroxybutyrate and selected for growth on nitrogen limiting minimal medium containing sucrose as a carbon source. Several cosmids were identified which apparently contain overlapping regions of the *R. meliloti* genome and the locus or loci responsible for the improved growth was mapped to a 2.5 kb region present in all of the cosmids identified in the screen. Further analysis revealed that the locus responsible for the growth improvement (*adhA*) is the first alcohol dehydrogenase gene isolated from the Rhizobiaceae. *R. meliloti adhA* is expressed in *A. eutrophus* under conditions where the native alcohol dehydrogenase activity is not observed. A model is proposed to explain these findings.

INTRODUCTION

Sucrose is the main photosynthate transported to nitrogen fixing root nodules such as those induced on *Medicago sativa* (alfalfa) roots by its symbiont, the soil bacterium *Rhizobium meliloti*. Our interest was drawn to the possibility that genes involved in sucrose metabolism may be required for or induced during the symbiotic process. When nodulated leguminous plants are provided $^{14}CO_2$ for photosynthesis, they incorporate the label and transport radiolabeled compounds to the root nodules (Antoniw and Sprent, 1978; Reibach and Streeter, 1983). Sucrose is the predominant photoassimilate transported in the phloem of most higher plants (Giaquinta, 1983). The primary radiolabeled compound detected in nodule and bacteroid fractions when nodulated lupin plants fix $^{14}CO_2$ is sucrose (Romanov et al., 1985). The detection of high concentrations of sucrose in the bacteroid fraction suggests sucrose is being transported across the symbiosome membrane and into the bacteroids. Transport of sucrose in the infection thread has not been investigated.

Uptake and hydrolysis of sucrose have been detected biochemically in free living *R. meliloti*. Martinez-de Drets et al. detected invertase activity in each of the 14 *R. meliloti* strains they tested (Martinez-de Drets et al., 1974). They also examined 64 strains of fast-and slow-growing rhizobia for evidence of the phosphoenol-pyruvate:sugar phosphotransferase system (PTS). The PTS is one of the major systems used by *E. coli* for importing sugars (Postma et al., 1993; Saier and Reizer, 1994). The first step in the PTS is transport of the sugar across the inner membrane coupled with phosphorylation of the substrate. No sucrose phosphorylase activity was detected in any of the strains tested by Martinez-de Drets et al., and they concluded that rhizobia do not use the PTS for uptake and hydrolysis of sucrose (Martinez-de Drets et al., 1974).

Glenn and Dilworth showed biochemically that *R. meliloti* SU47, the parent of Rm1021, has invertase and sucrose uptake activities, both of which are inducible by sucrose (Glenn and Dilworth, 1981). They demonstrated that disaccharide uptake is an active process in *R. meliloti*, which is abolished by inhibitors of protein synthesis and disrupters of proton motive force. No *R. meliloti* mutants have been reported which fail to grow on sucrose while retaining the ability to grow on glucose, one predicted phenotype of mutants defective in sucrose transport or hydrolysis.

The original aim of this study was to identify *R. meliloti* genes involved in sucrose uptake and metabolism. We first explored the possibility of using the indicator plates EMB and Maconkey in a screen for *R. meliloti* mutants unable to utilize disaccharides, but found that the standard formulations of these media used to detect sugar fermentation in *E. coli* are not useful for the detection of disaccharide utilization by *R. meliloti*. We therefore chose a different strategy which involved screening a *R. meliloti* genomic library for cosmids which permit a heterologous strain to grow on sucrose. However, our preliminary experiments did not yield any cosmids which allow an *Escherichia coli* strain to grow on sucrose. We then selected the gram-negative soil organism *Alcaligenes eutrophus* as a host in which to screen for *R. meliloti* genes involved in sucrose metabolism. *A. eutrophus* and *R. meliloti* both have high G+C content, and we knew from previous work that it is possible to express at least some *R. meliloti* genes in *A. eutrophus* (Chapter 2). The screen was conducted under limiting nitrogen to mimic conditions in the infection thread and pre-nitrogen fixation phase of symbiosis. An *A. eutrophus phbC* mutant, which does not produce poly-β-hydroxybutyrate (PHB) was used as the host. Use of this strain facilitated screening on sucrose and fructose plates because it eliminated PHB-related size and morphology differences seen when wild type cells are grown on the selective medium.

Characterization of cosmids which improved growth of *A. eutrophus* on nitrogen limiting sucrose medium led to the unexpected identification of an *R. meliloti* homologue of microbial alcohol dehydrogenase (*adh*) genes. The *adhA* gene is the first alcohol dehydrogenase gene isolated from the Rhizobiaceae. Further studies showed the *R. meliloti adhA* gene is expressed in the heterologous strain under the experimental conditions and suggest a model for the observed results.

MATERIALS AND METHODS

Strains and growth media

Strains and plasmids used in this paper are listed in Table 3-1. Bacterial strains were routinely grown in LB medium (Maniatis et al., 1982), which was supplemented with 2.5 mM $MgSO_4$ and 2.5 mM $CaCl_2$ in the case of *R. meliloti*. The defined medium MM1 (Peoples and Sinskey, 1989b), containing 0.05% $(NH_4)_2SO_4$ and 0.5% (w/v) fructose or 0.4% (w/v) filter sterilized sucrose, was used to assay growth of *A. eutrophus* strains. The nitrogen rich medium MM1N uses the same formulation as MM1 with the exception that the concentration of $(NH_4)_2SO_4$ is increased to 0.2%. Antibiotics were used at the following concentrations: ampicillin (Amp), 150 μg/ml; chloramphenicol (Cm), 20 μg/ml; gentamicin sulfate (Gm), 5 μg/ml for *E. coli*, 50 μg/ml for *R. meliloti*; kanamycin sulfate (Km), 50 μg/ml; nalidixic acid (Nal), 50 μg/ml; neomycin sulfate (Nm), 200 μg/ml; tetracycline (Tc), 10 μg/ml.

Genetic techniques

Conjugal transfer of plasmids was accomplished in triparental matings using pRK600 to provide transfer functions. Plasmid-borne insertions were recombined into the *R. meliloti* genome via homogenotization as described, (Glazebrook and Walker, 1991) using pPH1JI or pR751 as the incompatible IncP plasmid. Insertions were then transduced using bacteriophage ϕM12 into strain Rm1021 to ensure a clean genetic background. Southern hybridization was performed to check the construction of each strain.

DNA manipulations

Plasmid and cosmid DNA was isolated from overnight cultures of *E. coli* by the alkaline lysis method (Maniatis et al., 1982) or by purification over a Qiagen column. DNA modifying enzymes were used according to the instructions of the supplier (New England Biolabs, Beverly, MA or Takara, Japan). Gene Screen Plus membranes (Dupont/NEN, Boston, MA) were used for Southern hybridization. Radiolabeled DNA probes were prepared with the NEBlot random labeling kit (New England Biolabs, Beverly, MA) and ^{32}P-α-dCTP from Dupont/NEN (Boston, MA) or Amersham.

DNA Sequencing and analysis

The inserts of plasmids pLW121, pLW122 and pLW146 were sequenced from one or both ends. Double stranded DNA template was subjected to fluorescently labeled dideoxy termination reactions and run on an ABI Prism sequencing apparatus at the MIT Biopolymers Laboratory. Contigs were prepared with Assembly Lign software (Kodak/IBI). Database searches were performed using the BLAST programs (Altschul et al., 1990; Gish and States, 1993) to search databases maintained by the National Center for Biotechnology Information. Alignments of peptide sequences were performed using the GCG software package (Genetics Computer Group, 1991). The DNA sequence reported in this chapter has been deposited in GenBank and assigned the accession AF031940.

Isolation of cosmids which improve the growth of A. eutrophus phbC *on sucrose*

An *R. meliloti* genomic library in pLAFR1 (Friedman et al., 1982) was mated into *A. eutrophus* PHB#2 (Peoples and Sinskey, 1989b), and transconjugants were selected on MM1 plates containing 0.05% $(NH_4)_2SO_4$, tetracycline, and 0.4% sucrose

as the sole carbon source. Strains carrying pLW120-1 through pLW120-8 produced visible colonies after 4 days of incubation which grew to 3-4 mm in diameter after 11 days of incubation. A control strain carrying the vector pLAFR1 did not produce visible colonies even after 11 days of incubation at 30°C, although translucent microcolonies could be observed under magnification. All strains tested produced 4 mm colonies within 4 days on medium containing the permissive carbon source fructose.

Construction of disruptions in the R. meliloti adhA *gene and downstream region*

The kanamycin/neomycin resistance cassette from miniTn*5*-Km (de Lorenzo et al., 1990) was cloned into the *Bgl*II site of pLW140 to construct pLW141. pLW141 was digested with *Kpn*I and *Bam*HI and the insert was cloned into pSW213 to construct pLW142. In parallel, the kanamycin/neomycin resistance cassette from miniTn*5*-Km (de Lorenzo et al., 1990) was cloned into the *Hin*dIII site of pLW123 and the *Bam*HI site of pLW123 to construct plasmids pLW127 and pLW129, respectively. pLW142, pLW127 and pLW129 were each mated into Rm1021 and the drug marker was recombined into the genome. To construct strains Rm9612, Rm9610, and Rm9611, ϕM12 was employed to transduce the homogenotized antibiotic resistance marker into Rm1021 to ensure a clean genetic background. The location of each insertion was confirmed by Southern hybridization.

Mapping the adhA *locus*

The *adhA* locus was mapped to one of the three *R. meliloti* replicons using the method of Finan et al. (Finan et al., 1986). Genomic DNA from Rm1021 and *Agrobacterium tumefaciens* strains At123, At125 and At128 was digested with *Eco*RI, subjected to electrophoresis in a 0.6% agarose gel, transferred to a Gene Screen Plus

membrane and probed with the insert from pLW146, which contains the entire *R. meliloti adhA* gene. Additional mapping was performed by transduction of the Km^R/Nm^R marker from Rm9612 into the pRmeSU47a deletion strains GMI255, GMI766, GMI956, GMI963 and their parent RCR2011.

Assays for alcohol dehydrogenase activity in vivo

Alcohol dehydrogenase activity was detected *in vivo* on solid medium using Tetrazolium plates. The medium was formulated as described and supplemented with 0.5% ethanol and 0.001% 2,3,5-triphenyl tetrazolium chloride (Dowds et al., 1988). Bacterial strains were scored as positive for alcohol dehydrogenase activity if they formed dark red colonies on the Tetrazolium indicator plates but unpigmented colonies on Tetrazolium plates made without the ethanol substrate.

Detection of alcohol dehydrogenase activity in electrophoretic gels

Crude lysates were prepared from bacterial strains which were grown to saturation in LB medium. Cells were pelleted and resuspended in 1 ml of 0.1 M phosphate buffer pH 6.0 for every 50 ml of culture. Cells were lysed by sonication, and the cellular debris was removed by centrifugation in a tabletop microcentrifuge. Protein concentration was determined by the method of Bradford (Bradford, 1976) using a commercially available reagent (Bio-Rad) and bovine serum albumin (New England Biolabs) as a standard, and approximately 1 mg of protein was loaded onto 7.5% non-denaturing polyacrylamide gels. Gels were run at 4° C, typically at 12 mA until the dye front reached the bottom of the gel. Staining for alcohol dehydrogenase activity was accomplished by flooding a gel in a solution containing 50 ml 0.05 M Tris-HCl pH 8.5, 2 ml ethanol, 40 mg NAD, 10 mg Nitro Blue Tetrazolium (NBT) and 1 mg phenazine methosulfate (PMS) (Manchenko, 1994). The gel was then incubated at

37°C in the dark until dark bands appeared, rinsed in water and photographed. Bands which were present in gels incubated with alcohol but absent in gels incubated without alcohol were deemed to represent alcohol dehydrogenase activity.

Plant inoculation assays

Medicago sativa cv. Iroquois was obtained from Agway (Plymouth, IN). *R. meliloti* strains were tested for the ability to nodulate alfalfa on nitrogen-free Jensen's medium as described (Leigh et al., 1985). Plants were grown in a constant temperature room at 21°C with a 14 hour light cycle. Observations were made weekly for at least six weeks. The presence of pink, cylindrical nodules on dark green healthy plants was taken as evidence that nitrogen fixation was occurring. Plants with ineffective nodules were stunted and chlorotic.

RESULTS

Isolation and characterization of cosmids carrying R. meliloti *DNA that confer improved growth of* A. eutrophus *on sucrose medium.*

We used *Alcaligenes eutrophus*, which is unable to utilize sucrose, as a tool to identify *R. meliloti* genes encoding sucrose utilization functions. An *R. meliloti* genomic library was mated into *A. eutrophus phbC* and transconjugants were selected on nitrogen limiting minimal medium (MM1) containing sucrose as the sole carbon source. Transconjugants were compared with an isogenic strain harboring pLAFR1 and scored for growth on the selective medium after four days. Fewer than 0.1% of transconjugants grew on the sucrose plates. Twenty-six cosmids isolated from three separate matings were selected for further study.

Each of these cosmids was mated into *E. coli* C2110 and then mated back into *A. eutrophus phbC*. Transconjugants were selected on the MM1 containing fructose and subsequently tested for their ability to grow on MM1 sucrose. Although all of the cosmids identified in the screen conferred a marked growth improvement to *A. eutrophus phbC* which was not seen in the strain harboring pLAFR1, none of the cosmids identified in this screen permitted *A. eutrophus phbC* to grow as well on sucrose as it does on fructose, the preferred carbon source. Growth was judged by colony size on solid medium. When grown on MM1 fructose plates, *A. eutrophus phbC* harboring pLAFR1 or any of the cosmids identified in the screen formed colonies that grew to 4 mm in diameter after four days of incubation. Because *A. eutrophus* cannot utilize sucrose, *A. eutrophus phbC* harboring pLAFR1 does not form visible colonies on MM1 sucrose medium. However, because *A. eutrophus* is a facultative chemolithotroph capable of using CO_2 as a carbon source, *A. eutrophus phbC* carrying pLAFR1 forms microcolonies on MM1 sucrose medium which are translucent and can only be seen upon examination of the plates

with a dissecting microscope. In contrast, *A. eutrophus phbC* carrying any of the cosmids identified in this screen produced colonies which were clearly visible after four days of incubation. These colonies would reach 3-4 mm in diameter if incubation was continued for eleven days.

The cosmids identified in this screen each contain approximately 20 kb of *R. meliloti* genomic DNA cloned into the *Eco*RI site of pLAFR1. Analysis of *Eco*RI restriction digests showed that the 26 cosmids can be classified into eight different restriction patterns. One cosmid from each of the eight restriction patterns was chosen for further study, and these eight cosmids were designated pLW120–1 through pLW120–8. The complete physical map of the insert of pLW120-1 is shown in Figure 3-1A. The eight cosmids appear to contain overlapping regions of the *R. meliloti* genome because some restriction fragments are present in more than one cosmid. Strikingly, two bands are present in each of the eight cosmids: 4.2 kb and 1.7 kb *Eco*RI fragments.

The locus (loci) responsible for improved growth on sucrose maps to a 2.5 kb Eco*RI-*Hin*dIII region.*

To identify which locus or loci were responsible for this phenotype, we tested DNA subcloned from pLW120-1 for its ability to promote growth on sucrose. Since it seemed likely the locus or loci would be found in the region conserved in all eight cosmid patterns, the 4.2 and 1.7 kb *Eco*RI fragments from pLW120-1 were cloned into pBluescript SK+ and subsequently subcloned into the broad host range vector pSW213 to produce pLW123 and pLW124, respectively. As shown in Figure 3-1A, neither the 4.2 kb nor the 1.7 kb *Eco*RI fragment is able to promote the growth of *A. eutrophus phbC* on limiting nitrogen minimal medium containing sucrose as the sole carbon source.

Restriction mapping and Southern hybridization (data not shown) revealed that these two *Eco*RI fragments are adjacent to each other and overlap with a 4.6 kb *Hin*dIII fragment from pLW120-2. The 4.6 kb *Hin*dIII fragment from pLW120-2 was subcloned into pSW213 to produce pLW147. pLW147 was found to be able to confer improved growth to *A. eutrophus phbC* under the experimental conditions. Based on these observations, it seemed most likely that the locus (or loci) responsible for the phenotype is located in the 2.5 kb region between the 5′ *Eco*RI site of pLW124 and the 3′ *Hin*dIII site of pLW147.

Identification of an open reading frame encoding alcohol dehydrogenase.

The 2.5 kb DNA sequence shown in Figure 3-2 contains two large open reading frames. The first open reading frame consisted of 678 nucleotides and encoded a deduced peptide of 226 amino acid residues. This deduced peptide shows only weak homology to a hypothetical protein translated from an ORF of *Mycobacterium tuberculosis* (Philipp et al., 1996) and we named the locus *orf1*. This locus is entirely contained within pLW124, which is unable to promote improved growth on sucrose. Therefore, this ORF was either unrelated to, or was not sufficient for, the observed phenotype.

The second open reading frame overlaps the junction between the 1.7 kb and 4.2 kb *Eco*RI fragments. Because these two *Eco*RI fragments were present in all of the cosmids identified in this study, and neither is able to promote the growth advantage by itself, we focused our attentions on this open reading frame. This open reading frame contained 1023 nucleotides and encoded a deduced protein of 341 amino acids which was determined to have high homology with previously cloned microbial alcohol dehydrogenases, and we therefore named it *adhA*. Three closely related alcohol dehydrogenase proteins from *Bacillus stearothermophilus*

have the highest homology with AdhA (Sakoda and Imanaka, 1992; Cannio et al., 1994; Robinson et al., 1994). The deduced peptide AdhA is 74% similar and 56% identical to Adh-T from *B. stearothermophilus*. An alignment of AdhA and Adh-T is shown in Figure 3-3. *R. meliloti* AdhA deduced peptide shows homology with many previously studied alcohol dehydrogenase proteins, including those of *Emericella nidulans* and *Schizosaccharomyces pombe* (Russell and Hall, 1983). All of these proteins are classified as long-chain zinc-containing enzymes, because the deduced peptide chains contain approximately 350 amino acids, and some members of the class (Ecklund et al., 1976) have been shown experimentally to contain zinc ions.

Determination that the second open reading frame encoded alcohol dehydrogenase prompted additional experiments. Alcohol dehydrogenase is not predicted to be involved in sucrose transport or hydrolysis. To understand why this locus provided a growth advantage to *A. eutrophus phbC* on nitrogen limiting sucrose medium, we studied alcohol dehydrogenase activity in *A. eutrophus* in the presence or absence of heterologous *R. meliloti adhA*, and tested whether improved growth is observed under nitrogen rich conditions, or in a wild type *A. eutrophus* strain H16 background.

R. meliloti adhA *is expressed in* A. eutrophus *and has a different expression pattern than the native alcohol dehydrogenase.*

To examine alcohol dehydrogenase activity *in vivo*, strains were streaked on plates containing 2,3,5-triphenyl tetrazolium chloride and ethanol (Dowds et al., 1988). On this medium, Rm1021 produces colonies which are dark red from center to perimeter. Colonies produced by both *phbC* and *phbC*$^+$ *A. eutrophus* strains are white on the perimeter but contain dark red centers. In contrast, *A. eutrophus*

phbC or *phbC*$^+$ strains carrying pLW120-1 or pLW147 produce colonies which are red from center to perimeter, while strains carrying pLAFR1 have the same morphology as strains without a plasmid. This test shows that *R. meliloti adhA* is expressed in *A. eutrophus* and has a different pattern of expression than the native *A. eutrophus* alcohol dehydrogenase. These results are consistent with constitutive expression of *R. meliloti adhA* in *A. eutrophus*, and expression of *A. eutrophus* native alcohol dehydrogenase under low oxygen or anaerobic conditions.

R. meliloti adhA *offers no growth advantage on sucrose to* phbC$^+$ A. eutrophus

To determine the conditions under which pLW120-1 improves the growth of *A. eutrophus*, we examined the effect of changing the genetic background or nitrogen content of the medium. *A. eutrophus* strains carrying pLAFR1, pLW120-1 and pLW147 were constructed in both *phbC* and *phbC*$^+$ backgrounds. These six strains were streaked on MM1 and the nitrogen rich medium MM1N containing sucrose as the sole carbon source. Plates were incubated at 30°C for four to seven days and scored for growth. As expected, pLAFR1 conferred no growth advantage in either genetic background. pLW120-1 and pLW147 improved the growth of *A. eutrophus phbC* on sucrose plates regardless of the concentration of nitrogen, but provided no advantage to the *phbC*$^+$ strain. These data show that alcohol dehydrogenase offers a growth advantage only in the PHB$^-$ background.

R. meliloti adhA *does not permit growth on sucrose in highly aerated cultures*

In order to determine the doubling time of *A. eutrophus phbC* strains harboring pLW120-1 or pLAFR1 we monitored the cell density of liquid cultures grown at 30°C on a roller drum. We inoculated MM1 sucrose Tc and MM1 fructose Tc liquid cultures with washed cells of *A. eutrophus phbC* carrying pLAFR1 or pLW120-1 and

measured optical density at defined time points using a Klett meter. While the MM1 fructose Tc cultures grew well and had a doubling time of approximately 4 hours, the cultures grown in MM1 sucrose Tc showed no growth even after 3 or more days of incubation. This result shows that *R. meliloti adhA* does not promote growth on sucrose in aerated cultures.

Strains carrying disruptions in the R. meliloti *genome in and near* adhA *retain alcohol dehydrogenase activity and are Fix*$^+$

In order to examine the phenotype of a *R. meliloti adh* mutant, we constructed disruptions within and downstream of *adhA* as illustrated in Figure 3-1B and Figure 3-2, and then recombined these disruptions into the *R. meliloti* genome. Colonies produced by these *R. meliloti* strains on Tetrazolium assay plates appear identical to those produced by the wild type strain. To test whether *adhA* was required for symbiosis, alfalfa plants were inoculated with Rm1021 derivatives containing insertions either in the *adhA* gene or downstream of it. Like plants inoculated with Rm1021, all of the plants inoculated with Rm1021 derivatives had dark green leaves and pink nodules 28 days after inoculation, indicating that *adhA* function is not required for the development of an effective symbiosis.

Detection of AdhA protein in polyacrylamide gels

Alcohol dehydrogenase activity in wild type and *adhA* strains of *R. meliloti* and *A. eutrophus phbC* strains harboring pLAFR1 or pLW120-1 was examined biochemically by staining proteins isolated from saturated cultures for alcohol dehydrogenase activity in non-denaturing polyacrylamide gels (Figure 3-4). The most prominent band of alcohol dehydrogenase activity seen in wild type *R. meliloti* is missing in the lane containing protein from the *adhA* mutant,

strongly suggesting that this protein is encoded by the *adhA* gene. Significantly, a very prominent protein band of the same mobility is seen in the lane containing protein isolated from *A. eutrophus* harboring pLW120-1, but not an isogenic strain carrying the vector pLAFR1. This observation argues strongly that the prominent band is indeed encoded by *adhA*. In addition to AdhA, the protein isolated from Rm1021 appears to contain two proteins with strong alcohol dehydrogenase activity, and several minor bands. The additional bands of NAD-dependent ethanol dehydrogenase activity suggest the existence of additional *R. meliloti adh* genes.

adhA *maps to pRmeSU47a, the smaller of two megaplasmids*

R. meliloti has three replicons, the chromosome and two megaplasmids of 1.4 and 1.7 Mb (Sobral et al., 1991). The *adhA* locus was mapped by Southern hybridization using the method of Finan et al. (Finan et al., 1986). *Eco*RI-digested genomic DNA from Rm1021 and *Agrobacterium tumefaciens* strains cured of the Ti plasmid and carrying (a) *R. meliloti* megaplasmid pRmeSU47a, (b) *R. meliloti* megaplasmid pRmeSU47b or (c) no megaplasmid was probed with radiolabeled pLW123 DNA, which contains the C-terminal half of *adhA*. Strongly hybridizing bands of 4.2 kb were observed in DNA isolated from Rm1021 and the *A. tumefaciens* strain carrying pRmeSU47a, but not the other two *A. tumefaciens* strains, indicating that the locus maps to pRmeSU47a. The fact that the *A. tumefaciens* strain cured of its Ti plasmid does not hybridize to pLW123 insert under high stringency conditions shows that this strain does not have a close homologue of *R. meliloti adhA*, despite the fact that *R. meliloti* and *A. tumefaciens* are fairly closely related in evolutionary terms.

Additional mapping was performed by transduction. To determine whether the *adhA* locus maps near the well characterized *nod* region of pRmeSU47a, the kanamycin/neomycin resistance marker from Rm9612 was transduced into

R. meliloti strains carrying large deletions of the *nod-nif* region. The Km^R/Nm^R marker was successfully transduced into RCR2011 and the deletion strains GMI255, GMI766, GMI956, and GMI963, indicating that the *adhA* locus does not lie within the regions deleted in these strains.

DISCUSSION

Several classes of genes could have caused increased growth of *A. eutrophus* on plates containing sucrose as a sole carbon source. Our original model suggested genes involved in sucrose transport or hydrolysis would be identified in this screen. Instead, the cosmids that were identified in a screen for loci that increase the growth of *A. eutrophus* on nitrogen limited minimal medium with sucrose as a carbon source encoded a *R. meliloti* homologue of long chain, zinc-containing alcohol dehydrogenase (Adh) proteins.

R. meliloti possesses *multiple Adh proteins*

Our *in vivo* results from Tetrazolium assay plates show that an *R. meliloti adhA* mutant retains alcohol dehydrogenase activity, and our *in vitro* results show that *R. meliloti* produces three prominent proteins with alcohol dehydrogenase activity. This suggests the presence of at least one additional *adh* gene. Bacteria and other microbes often contain more than one gene encoding alcohol dehydrogenase (Reid and Fewson, 1994). Our results are consistent with the report of Rigaud and Trinchant, who visualized three isozymes of *R. meliloti* alcohol dehydrogenase after starch electrophoresis (Rigaud and Trinchant, 1973). In mammals and plants, different isoforms of Adh can have tissue-specific expression patterns. The low O_2 concentration in the nodule, which is required for active nitrogenase, causes stress to the plant cells in the nodule. Consequently, enzymes involved in fermentative metabolism, including alcohol dehydrogenase, are stimulated. Alfalfa Adh activity has been isolated and characterized, and is elevated in nodules compared to root tissue (Irigoyen et al., 1992). It is possible that *R. meliloti* possesses *adh* genes which are differentially expressed in the free-living and bacteroid states. *R. meliloti* Adh activity has been detected in bacteroids and is elevated compared to the free-living

state (Irigoyen et al., 1992), but nothing is know about whether different genes are responsible for Adh activity in free-living and microsymbiotic cells. Because of these observations we were interested in whether *adhA* was involved in symbiosis. However, *R. meliloti* strains carrying disruptions in the *adhA* gene are still able to establish an effective symbiosis with alfalfa.

A model for the role of R. meliloti adhA *in* A. eutrophus *metabolism*

It is interesting that although *R. meliloti adhA* is not involved directly in sucrose metabolism, we isolated it because it provides a growth advantage to a *A. eutrophus phbC* strain grown on minimal medium under conditions of limiting nitrogen. These conditions have been found to favor PHB production in wild type *A. eutrophus* strains and to result in excretion of pyruvate into the medium in strains deficient in biosynthesis of PHB (Steinbüchel and Schlegel, 1989). This results in the acidification of the medium, which is detrimental to growth (Steinbüchel and Schlegel, 1989).

Our tetrazolium assays demonstrate that *adhA* is expressed constitutively in *A. eutrophus*, whereas *A. eutrophus* alcohol dehydrogenase activity is seen only in the centers of colonies, consistent with the report that the native *A. eutrophus* alcohol dehydrogenase is expressed under conditions of limiting oxygen (Jendrossek et al., 1988). We propose that by converting acetaldehyde to ethanol, *R. meliloti* AdhA provides a physiological benefit to *A. eutrophus phbC* under the conditions tested by improving the conversion of pyruvate to acetaldehyde and thus deacidifying the medium. AdhA may also provide a benefit by recycling NADH back to NAD, permitting glycolysis to proceed in the anaerobic conditions within a colony. Figure 3-5 diagrams the relevant steps in intermediary metabolism. Because we could not grow *A. eutrophus* strains harboring pLW120-1 in minimal

liquid cultures containing sucrose as the sole carbon source, it was not possible to assay for acidification of the medium nor accumulation of metabolic intermediates.

Wild type *A. eutrophus* grown under limiting nitrogen conditions does not accumulate pyruvate and acetaldehyde, but instead proceeds with PHB synthesis. Since there is no acidification of the medium, and NAD does not need to be recycled, our model predicts that *R. meliloti adhA* will offer no growth advantage to wild type *A. eutrophus*. This prediction was borne out by our experimental findings.

No *R. meliloti* mutants defective in sucrose transport or hydrolysis have been reported, which may indicate that there are redundant genes for this capacity, or that such genes are absolutely required for growth. Although our initial screening conditions produced unexpected results, it seemed that the approach could be useful for identifying genes involved in sucrose metabolism. In subsequent experiments (Chapter 4), we modified the screen and identified loci involved in *R. meliloti* α-glucoside metabolism.

Table 3-1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
Strain		
Rhizobium meliloti		
Rm1021	SU47 *str-21*	(Meade et al., 1982)
Rm9610	Rm1021 with Km^R insertion 127	this work
Rm9611	Rm1021 with Km^R insertion 129	this work
Rm9612	Rm1021 *adhA142*	this work
RCR2011	RCR2011 ≡ SU47	(Truchet et al., 1985)
GMI255	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI766	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI956	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
GMI963	RCR2011 derivative with pRmeSU47a deletion	(Truchet et al., 1985)
Alcaligenes eutrophus		
H16	wild type, Sm^S	(Peoples and Sinskey, 1989b)
PHB#2	H16 *phbC*::Tn*5*	(Peoples and Sinskey, 1989b)
Agrobacterium tumefaciens		
At123	GMI9023 ≡ GMI9050 cured of pAtC58 and pTi	(Rosenberg and Huguet, 1984)
At125	GMI9023 pRmeSU47bΩ5007::Tn*5-oriT*	(Finan et al., 1986)
At128	GMI9023 pRmeSU47aΩ30::Tn*5-11*	(Finan et al., 1986)
Escherichia coli		
DH5α	*supE44 ΔlacU169(ϕ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Clontech
Plasmid		
pRK600	pRK2013 *npt*::Tn*9*, Cm^R	(Finan et al., 1986)
pPH1JI	IncP, Gm^R, Sp^R, Cm^R	(Beringer et al., 1978)
pR751	IncP, Tp^R	(Meyer and Shapiro, 1980)
pLAFR1	IncP cloning vector, Tc^R	(Friedman et al., 1982)
pSW213	IncP broad host range vector, Tc^R	(Chen and Winans, 1991)
pBluescript SK+	Derivative of pUC19 with f1(+) *oriR*	Stratagene
pBluescript II KS+	Derivative of pUC19 with f1(+) *oriR*	Stratagene
pUTminiTn*5*Km	Amp^R, Km^R/Nm^R; source of Km^R cassette	(de Lorenzo et al., 1990)
pLW120-1	pLAFR1 clone carrying *R. meliloti adhA*	this work
pLW120-2	pLAFR1 clone carrying *R. meliloti adhA*	this work
pLW121	pBluescript SK+ with 4.2kb *Eco*RI fragment from pLW120-1	this work
pLW122	pBluescript SK+ with 1.7kb *Eco*RI fragment from pLW120-1	this work
pLW123	pSW213 with 4.2 kb *Eco*RI from pLW121	this work
pLW124	pSW213 with 1.7 kb *Eco*RI from pLW122	this work
pLW127	Km^R cassette in *Hin*dIII site of pLW123	this work
pLW129	Km^R cassette in *Bam*HI site of pLW123	this work
pLW140	pBluescript II KS+ with 0.9 kb *Eco*RI-*Hin*dIII fragment from pLW121	this work
pLW141	*Bam*HI-cut Km^R cassette in *Bgl*II site of pLW140	this work
pLW142	pSW213 with 3.1 kb *Bam*HI-*Kpn*I fragment from pLW141	this work
pLW146	pBluescript KSII+ with 4.6 kb *Hin*dIII fragment from pLW120-2	this work
pLW147	pSW213 with 4.6 kb *Hin*dIII from pLW146	this work

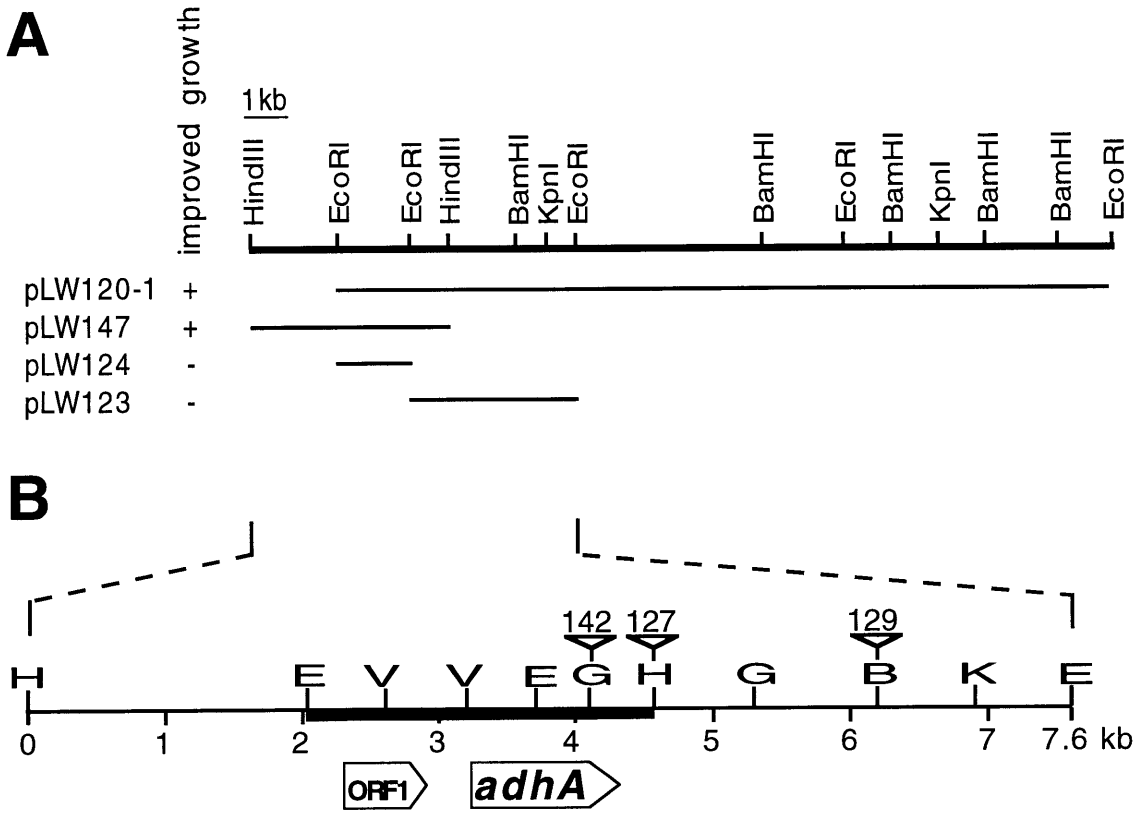
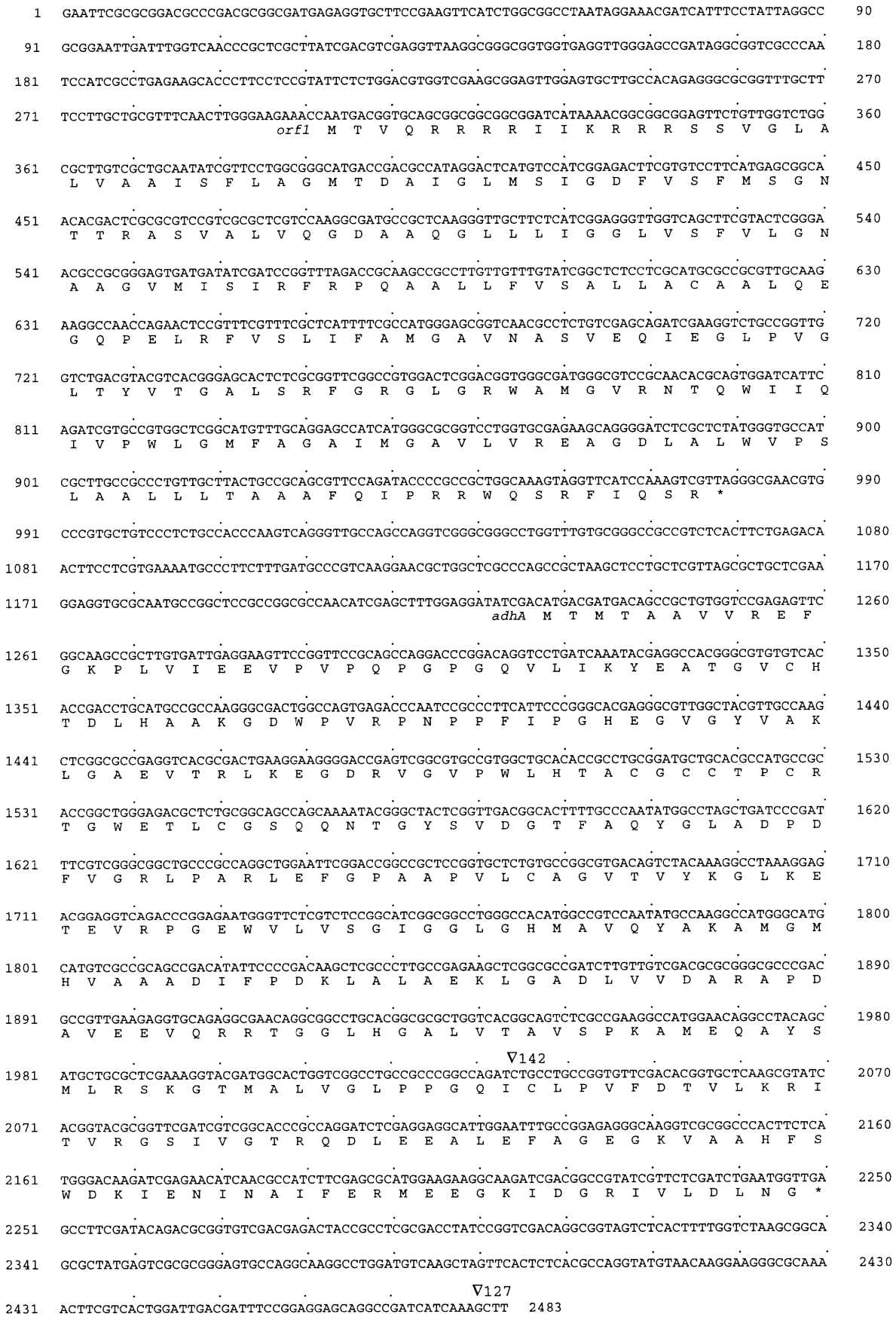


Figure 3-1.

(A) Restriction map of *R. meliloti* DNA present in pLW120-1, its derivatives and pLW147. Horizontal lines below the restriction map represent sequences remaining in subclones. The ability of a plasmid to promote growth of *A. eutrophus phbC* on MM1 sucrose is indicated to the right of the plasmid name.

(B) Higher resolution restriction map of DNA from the subclones pLW147, pLW124 and pLW123. Vertical lines topped by triangles indicate the positions of antibiotic resistance cassette insertions. The thick horizontal line marks the region that was sequenced. The positions of *orf1* and *adhA* are shown. The positions of *Hin*dIII (H), *Eco*RI (E), *Eco*RV (V), *Bgl*II (G), *Bam*HI (B) and *Kpn*I (K) restriction sites are shown.

Figure 3-2.
Nucleotide sequence of the *Hin*dIII-*Eco*RI region indicated in Figure 3-1B and deduced amino acid sequence of *orf1* and *adhA*. Downward arrows indicate the sites of Km^R/Nm^R insertions described in the text and in Figure 3-1B.


1 GAATTCGCGCGGACGCCCGACGCGGCGATGAGAGGTGCTTCCGAAGTTCATCTGGCGGCCTAATAGGAAACGATCATTTCCTATTAGGCC 90
91 GCGGAATTGATTTGGTCAACCCGCTCGCTTATCGACGTCGAGGTTAAGGCGGGCGGTGGTGAGGTTGGGAGCCGATAGGCGGTCGCCCAA 180
181 TCCATCGCCTGAGAAGCACCCTTCCTCCGTATTCTCTGGACGTGGTCGAAGCGGAGTTGGAGTGCTTGCCACAGAGGGCGCGGTTTGCTT 270
271 TCCTTGCTGCGTTTCAACTTGGGAAGAAACCAATGACGGTGCAGCGGCGGCGGCGGATCATAAAACGGCGGCGGAGTTCTGTTGGTCTGG 360
orf1 M T V Q R R R R I I K R R R S S V G L A
361 CGCTTGTCGCTGCAATATCGTTCCTGGCGGGCATGACCGACGCCATAGGACTCATGTCCATCGGAGACTTCGTGTCCTTCATGAGCGGCA 450
L V A A I S F L A G M T D A I G L M S I G D F V S F M S G N
451 ACACGACTCGCGCGTCCGTCGCGCTCGTCCAAGGCGATGCCGCTCAAGGGTTGCTTCTCATCGGAGGGTTGGTCAGCTTCGTACTCGGGA 540
T T R A S V A L V Q G D A A Q G L L L I G G L V S F V L G N
541 ACGCCGCGGGAGTGATGATATCGATCCGGTTTAGACCGCAAGCCGCCTTGTTGTTTGTATCGGCTCTCCTCGCATGCGCCGCGTTGCAAG 630
A A G V M I S I R F R P Q A A L L F V S A L L A C A A L Q E
631 AAGGCCAACCAGAACTCCGTTTCGTTTCGCTCATTTTCGCCATGGGAGCGGTCAACGCCTCTGTCGAGCAGATCGAAGGTCTGCCGGTTG 720
G Q P E L R F V S L I F A M G A V N A S V E Q I E G L P V G
721 GTCTGACGTACGTCACGGGAGCACTCTCGCGGTTCGGCCGTGGACTCGGACGGTGGGCGATGGGCGTCCGCAACACGCAGTGGATCATTC 810
L T Y V T G A L S R F G R G L G R W A M G V R N T Q W I I Q
811 AGATCGTGCCGTGGCTCGGCATGTTTGCAGGAGCCATCATGGGCGCGGTCCTGGTGCGAGAAGCAGGGGATCTCGCTCTATGGGTGCCAT 900
I V P W L G M F A G A I M G A V L V R E A G D L A L W V P S
901 CGCTTGCCGCCCTGTTGCTTACTGCCGCAGCGTTCCAGATACCCCGCCGCTGGCAAAGTAGGTTCATCCAAAGTCGTTAGGGCGAACGTG 990
L A A L L L T A A A F Q I P R R W Q S R F I Q S R *
991 CCCGTGCTGTCCCTCTGCCACCCAAGTCAGGGTTGCCAGCCAGGTCGGGCGGGCCTGGTTTGTGCGGGCCGCCGTCTCACTTCTGAGACA 1080
1081 ACTTCCTCGTGAAAATGCCCTTCTTTGATGCCCGTCAAGGAACGCTGGCTCGCCCAGCCGCTAAGCTCCTGCTCGTTAGCGCTGCTCGAA 1170
1171 GGAGGTGCGCAATGCCGGCTCCGCCGGCGCCAACATCGAGCTTTGGAGGATATCGACATGACGATGACAGCCGCTGTGGTCCGAGAGTTC 1260
adhA M T M T A A V V R E F
1261 GGCAAGCCGCTTGTGATTGAGGAAGTTCCGGTTCCGCAGCCAGGACCCGGACAGGTCCTGATCAAATACGAGGCCACGGGCGTGTGTCAC 1350
G K P L V I E E V P V P Q P G P G Q V L I K Y E A T G V C H
1351 ACCGACCTGCATGCCGCCAAGGGCGACTGGCCAGTGAGACCCAATCCGCCCTTCATTCCCGGGCACGAGGGCGTTGGCTACGTTGCCAAG 1440
T D L H A A K G D W P V R P N P P F I P G H E G V G Y V A K
1441 CTCGGCGCCGAGGTCACGCGACTGAAGGAAGGGGACCGAGTCGGCGTGCCGTGGCTGCACACCGCCTGCGGATGCTGCACGCCATGCCGC 1530
L G A E V T R L K E G D R V G V P W L H T A C G C C T P C R
1531 ACCGGCTGGGAGACGCTCTGCGGCAGCCAGCAAAATACGGGCTACTCGGTTGACGGCACTTTTGCCCAATATGGCCTAGCTGATCCCGAT 1620
T G W E T L C G S Q Q N T G Y S V D G T F A Q Y G L A D P D
1621 TTCGTCGGGCGGCTGCCCGCCAGGCTGGAATTCGGACCGGCCGCTCCGGTGCTCTGTGCCGGCGTGACAGTCTACAAAGGCCTAAAGGAG 1710
F V G R L P A R L E F G P A A P V L C A G V T V Y K G L K E
1711 ACGGAGGTCAGACCCGGAGAATGGGTTCTCGTCTCCGGCATCGGCGGCCTGGGCCACATGGCCGTCCAATATGCCAAGGCCATGGGCATG 1800
T E V R P G E W V L V S G I G G L G H M A V Q Y A K A M G M
1801 CATGTCGCCGCAGCCGACATATTCCCCGACAAGCTCGCCCTTGCCGAGAAGCTCGGCGCCGATCTTGTTGTCGACGCGCGGGCGCCCGAC 1890
H V A A A D I F P D K L A L A E K L G A D L V V D A R A P D
1891 GCCGTTGAAGAGGTGCAGAGGCGAACAGGCGGCCTGCACGGCGCGCTGGTCACGGCAGTCTCGCCGAAGGCCATGGAACAGGCCTACAGC 1980
A V E E V Q R R T G G L H G A L V T A V S P K A M E Q A Y S
∇142
1981 ATGCTGCGCTCGAAAGGTACGATGGCACTGGTCGGCCTGCCGCCCGGCCAGATCTGCCTGCCGGTGTTCGACACGGTGCTCAAGCGTATC 2070
M L R S K G T M A L V G L P P G Q I C L P V F D T V L K R I
2071 ACGGTACGCGGTTCGATCGTCGGCACCCGCCAGGATCTCGAGGAGGCATTGGAATTTGCCGGAGAGGGCAAGGTCGCGGCCCACTTCTCA 2160
T V R G S I V G T R Q D L E E A L E F A G E G K V A A H F S
2161 TGGGACAAGATCGAGAACATCAACGCCATCTTCGAGCGCATGGAAGAAGGCAAGATCGACGGCCGTATCGTTCTCGATCTGAATGGTTGA 2250
W D K I E N I N A I F E R M E E G K I D G R I V L D L N G *
2251 GCCTTCGATACAGACGCGGTGTCGACGAGACTACCGCCTCGCGACCTATCCGGTCGACAGGCGGTAGTCTCACTTTTGGTCTAAGCGGCA 2340
2341 GCGCTATGAGTCGCGCGGGAGTGCCAGGCAAGGCCTGGATGTCAAGCTAGTTCACTCTCACGCCAGGTATGTAACAAGGAAGGGCGCAAA 2430
∇127
2431 ACTTCGTCACTGGATTGACGATTTCCGGAGGAGCAGGCCGATCATCAAAGCTT 2483

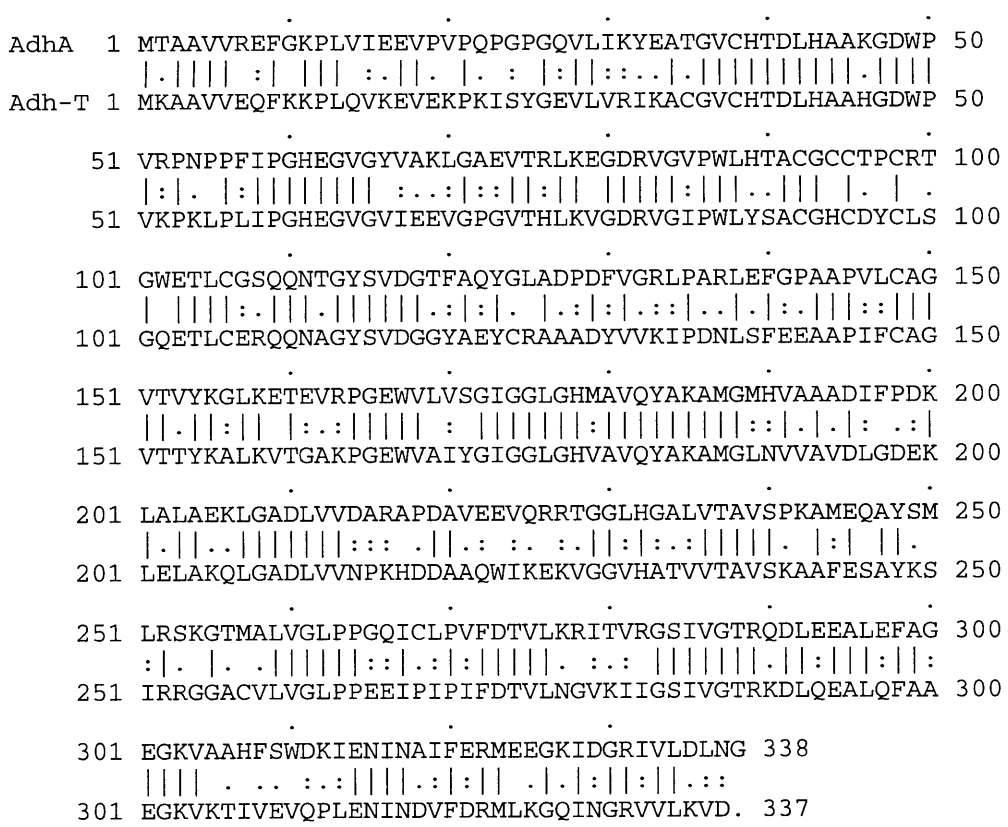


Figure 3-3.

Alignment of *R. meliloti* AdhA and alcohol dehydrogenase protein Adh-T from *Bacillus stearothermophilus*. Identical residues are indicated by a vertical line. Conserved substitutions are indicated by a dot or colon. The alignment was prepared using the GAP program of the Genetics Computing Group. Strongly conserved substitutions are indicated by a colon (comparison value ≥ 0.50) and weakly conserved substitutions are indicated by a dot (comparison ≥ value 0.10).

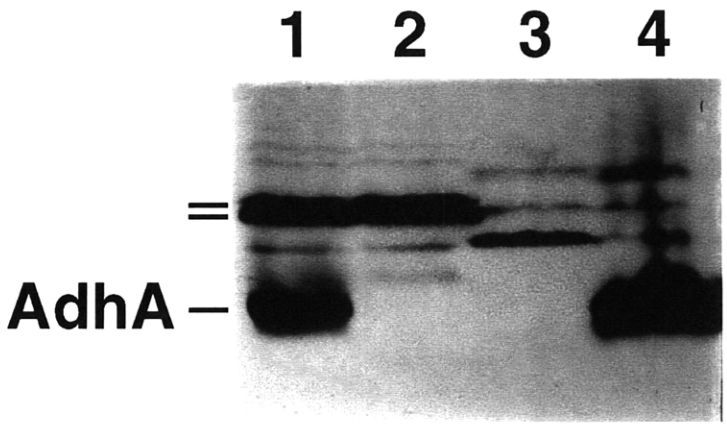


Figure 3-4

Alcohol dehydrogenase activity gel. Approximately 1 mg of protein from the indicated strains was separated on a 7.5% non-denaturing polyacrylamide gel and then stained for alcohol dehydrogenase activity. Lane 1: *Rhizobium meliloti* wild type strain Rm1021. Lane 2: *R. meliloti adhA* mutant strain Rm9612. Lane 3: *Alcaligenes eutrophus phbC* harboring pLAFR1. Lane 4: *A. eutrophus phbC* harboring pLW120-1. The three most prominent bands of alcohol dehydrogenase activity present in Lane 1 are indicated with horizontal lines.

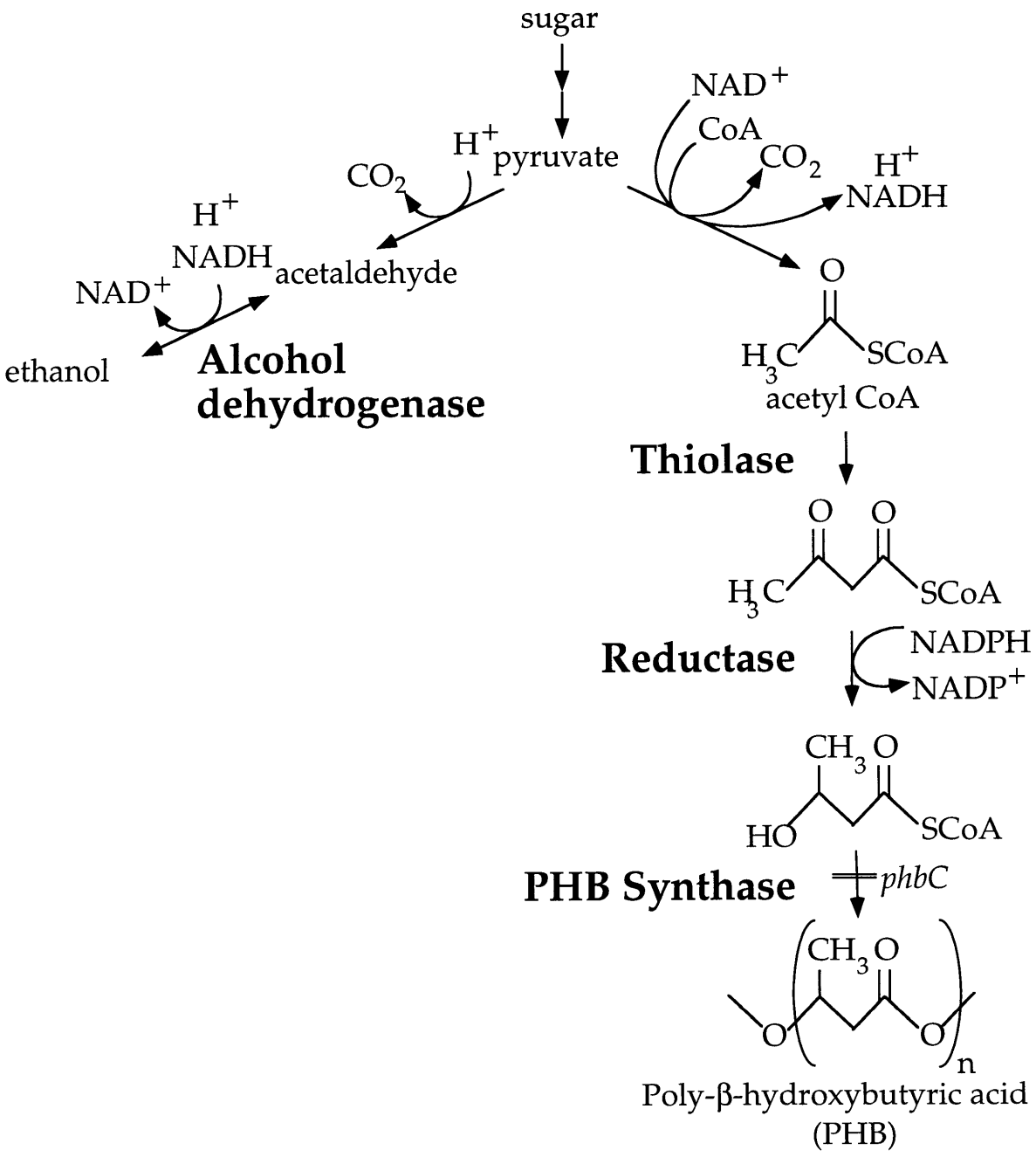


Figure 3-5.

Details of *A. eutrophus* metabolic pathways illustrating the proposed role of *R. meliloti adhA* in improving the growth of *A. eutrophus phbC* on nitrogen limiting sucrose medium. See text for details.

Chapter 4

A novel *Rhizobium meliloti* operon encodes an α-glucosidase and a periplasmic binding protein dependent transport system for α-glucosides

ABSTRACT

The most abundant carbon source transported into legume root nodules is photosynthetically produced sucrose, yet the importance of its metabolism by rhizobia *in planta* is not yet known. To identify genes involved in sucrose uptake and hydrolysis, we screened a *Rhizobium meliloti* genomic library and discovered a segment of *R. meliloti* DNA which allows *Alcaligenes eutrophus* to grow on the alpha-glucosides sucrose, maltose and trehalose. Tn*5* mutagenesis localized the required genes to a 6.8 kb region containing five open reading frames which were named *agl* for *a*lpha *gl*ucoside utilization. Four of these (*aglE*, *aglF*, *aglG* and *aglK*) appear to encode a periplasmic binding-protein dependent sugar transport system and one (*aglA*) appears to encode an alpha-glucosidase with homology to family 13 of glycosyl hydrolases. Cosmid-borne *agl* genes permit the uptake of radiolabeled sucrose into *A. eutrophus* cells. Analysis of the properties of *agl* mutants suggests that *R. meliloti* possesses at least one additional alpha-glucosidase as well as a lower affinity transport system for alpha-glucosides. It is possible that the Fix$^+$ phenotype of *agl* mutants on alfalfa is due to these additional functions. Loci found by DNA sequencing to be adjacent to *aglEFGAK* include a probable regulatory gene (*aglR*), *zwf* and *edd* which encode the first two enzymes of the Entner-Doudoroff pathway, *devB* which encodes a putative oxidoreductase, and a novel *Rhizobium*-specific repeat element.

INTRODUCTION

Photosynthetically derived sucrose is the main source of carbon to legume root nodules. In fact, sucrose is the first radiolabeled compound found in the root nodules and bacteroids of nodulated plants which are incorporating $^{14}CO_2$ via photosynthesis (Romanov et al., 1985). The identification of radiolabeled sucrose in the bacteroids in these studies suggests that sucrose is being transported across the symbiosome membrane.

Based on these observations, we thought it possible that sucrose uptake and hydrolysis may be required for, or induced during, symbiosis. However, the possible importance of sucrose metabolism during symbiosis has not yet been evaluated. To date, research on carbon metabolism in indeterminate nodules, such as those formed in the *Medicago sativa-Rhizobium meliloti* symbiosis, has emphasized the role of dicarboxylic acids in nitrogen fixation (Ronson et al., 1981; Bolton et al., 1986; Engelke et al., 1987; Yarosch et al., 1989). Although it has been demonstrated that transport of dicarboxylic acids is required for nitrogen fixation, it seems unlikely that dicarboxylic acids are required by bacteroids as a carbon source *per se*. *R. meliloti dct* mutants, which fail to transport dicarboxylic acids, cannot utilize carbon sources such as succinate in the free living state (Finan et al., 1988) and are Fix^- (Engelke et al., 1987; Yarosch et al., 1989). However, these strains are able to induce and invade nodules, and proceed through several stages of bacteroid development, becoming blocked just prior to active nitrogen fixation (Vasse et al., 1990). The efficiency of the TCA cycle is reduced in *dct* strains, and the defect in nitrogen fixation may be due to failure to produce enough ATP to power the nitrogenase holoenzyme.

Sucrose metabolism has been examined at the biochemical level in both fast- and slow-growing rhizobia (Martinez-de Drets et al., 1974; Glenn and Dilworth, 1981).

Studies of disaccharide metabolism demonstrated that sucrose hydrolysis and uptake activities are inducible in *R. meliloti*. The results of competition studies suggested that *R. meliloti* possesses at least three systems for disaccharide uptake: one system that transports sucrose, maltose and trehalose; a second which transports lactose; and a third which transports cellobiose (Glenn and Dilworth, 1981).

In these experiments, we sought to identify *R. meliloti* genes involved in sucrose transport or hydrolysis, so that we could begin to address the question of whether sucrose is utilized during nodule invasion or bacteroid development. Mutants of *R. meliloti* have been isolated which cannot utilize sucrose (Cerveñansky and Arias, 1984) or grow poorly on sucrose (Arias et al., 1979), but these strains fail to utilize several carbon sources and the defects in metabolism were found to be downstream of sucrose uptake or cleavage. No mutants of *Rhizobium* have been reported that are unable to utilize sucrose yet still able to utilize fructose or glucose, a key phenotype predicted for strains defective in sucrose hydrolysis or transport.

Since no *Rhizobium* mutants had been reported that were specifically defective in sucrose utilization, and we had not succeeded in isolating them by direct screening, we turned to a different strategy. This involved introducing a cosmid library of *R. meliloti* DNA into a heterologous host that could not utilize sucrose and selecting for derivatives that could grow on sucrose. This type of approach has been used successfully to identify sucrose utilization genes in the phosphoenolpyruvate-phosphotransferase system (PTS) by screening in *Escherichia coli* (Garcia, 1985). In our case, *E. coli* proved to be unsuitable, possibly because its G+C content is so much lower than *R. meliloti*, so instead we utilized *Alcaligenes eutrophus*, a gram-negative soil bacterium which has a high (~66%) G+C content and which we have shown in previous work to express the *phbC* gene of *R. meliloti* (Chapter 2).

We report here the identification of five *R. meliloti* genes which permit the growth of *A. eutrophus* on sucrose, maltose or trehalose. These genes evidently encode an α-glucosidase and a system for the transport of α-glucosides. A cosmid carrying these genes permits uptake of radiolabeled sucrose by *A. eutrophus*. In addition, we report the sequence of a putative regulatory locus (*aglR*), the Entner-Doudoroff genes of *R. meliloti*, a putative oxidoreductase gene, *devB*, and a novel *Rhizobium*-specific repeat element.

MATERIALS AND METHODS

Strains and growth media

Strains and plasmids used in this paper are listed in Table 4-1. Bacterial strains were routinely grown in LB medium (Maniatis et al., 1982), which was supplemented with 2.5 mM $MgSO_4$ and 2.5 mM $CaCl_2$ in the case of *R. meliloti*. The minimal medium M9 (Long et al., 1988) supplemented with 1 mM $MgSO_4$, 0.25 mM $CaCl_2$, 1 mg/ml D-biotin and 0.4% of filter-sterilized carbon source was used to assay the growth of *R. meliloti* strains. The defined medium MM1N [based on the medium MM1 (Peoples and Sinskey, 1989b) with the exception that the concentration of $(NH_4)_2SO_4$ is increased to 0.2%] supplemented with 0.5% (w/v) fructose or 0.4% (w/v) filter sterilized sucrose was used to assay growth of *A. eutrophus* strains. Where noted, NH_4Cl was substituted for $(NH_4)_2SO_4$, keeping constant the final concentration of nitrogen in the medium. Antibiotics were used at the following concentrations: ampicillin (Amp), 150 μg/ml; chloramphenicol (Cm), 20 μg/ml; gentamicin sulfate (Gm), 5 μg/ml for *E. coli*, 50 μg/ml for *R. meliloti*; kanamycin sulfate (Km), 50 μg/ml; nalidixic acid (Nal), 50 μg/ml; neomycin sulfate (Nm), 200 μg/ml; tetracycline (Tc), 10 μg/ml. To select for Tn*5* or miniTn*5*Km, kanamycin was used with *E. coli* and *A. eutrophus* and neomycin with *R. meliloti*.

Genetic techniques

Conjugal transfer of plasmids was accomplished in triparental matings using pRK600 to provide transfer functions. Plasmid-borne insertions were recombined into the *R. meliloti* genome via homogenotization as described, (Glazebrook and Walker, 1991) using pPH1JI or pR751 as the incompatible IncP plasmid. Insertions were then transduced using bacteriophage ϕM12 into strain Rm1021 to ensure a clean genetic background. Southern hybridization was performed to check the

construction of each strain. To obtain Gm^R/Sp^R derivatives of *R. meliloti* Tn*5*-induced mutations, Tn*5* insertions were replaced with Tn*5-233* (De Vos et al., 1986) as previously described (Glazebrook and Walker, 1991).

DNA manipulations

Plasmid and cosmid DNA was isolated from overnight cultures of *E. coli* by the alkaline lysis method (Maniatis et al., 1982) or by purification over a Qiagen column. DNA modifying enzymes were used according to the instructions of the supplier (New England Biolabs, Beverly, MA or Takara, Japan). Gene Screen Plus membranes (Dupont/NEN, Boston, MA) were used for Southern hybridization. Radiolabeled DNA probes were prepared with the NEBlot random labeling kit (New England Biolabs, Beverly, MA) and ^{32}P α-dCTP from Dupont/NEN (Boston, MA) or Amersham.

DNA Sequencing and analysis

Plasmids were purified for sequencing using a Qiagen plasmid mini kit. The sequencing strategy was based on a detailed restriction map of pLW200. Each of the 10 *Eco*RI fragments of pLW200 was subcloned into pBluescript SK+. These plasmids and defined subfragments cloned into pBluescript SK+ or pBluescript II KS+ were subjected to fluorescently labeled dideoxy termination reactions at the MIT Biopolymers laboratory or in an MJ-Research thermal cycler. The sequencing reactions were then separated on an ABI-Prism apparatus at the MIT Biopolymers laboratory or at the Molecular Biology facility at Dartmouth. Contigs were prepared using the SeqMan software program (Lasergene). Comparisons of nucleotide sequences and translated nucleotide sequences with GenBank were performed by using the BLAST algorithms (Altschul et al., 1990; Gish and States, 1993) to search

the databases maintained by the National Center for Biotechnology Information. Searches of SWISSPROT using Profile Scan to find PROSITE (Bairoch, 1992) patterns were performed with the resources maintained by ExPasy in Switzerland. Additional analysis was performed using the package of software developed by the Genetics Computing Group (GCG) (Genetics Computer Group, 1991), DNA Strider version 1.2, and the DNASTAR programs by Lasergene. The DNA sequence reported in this chapter has been submitted to GenBank and assigned the accession AF045609.

Isolation of cosmids which improve the growth of A. eutrophus *on sucrose*

An *R. meliloti* genomic library in pLAFR1 (Friedman et al., 1982) was mated into *A. eutrophus* H16, and transconjugants were selected on MM1N plates containing sucrose as the sole carbon source. In the initial screen, half of the selective plates contained NH_4Cl as a nitrogen source and half contained $(NH_4)_2SO_4$, but all subsequent experiments used $(NH_4)_2SO_4$ as a nitrogen source. *A. eutrophus* strains carrying pLW200 and related plasmids produced visible colonies after 3 days of incubation and were scored for growth on sucrose after 5 days. After 11 days of incubation at 30° C, these strains produced colonies approximately 4 mm in diameter. A control strain carrying pLAFR1 did not produce visible colonies even after 11 days of incubation at 30° C, although translucent microcolonies could be observed under magnification. All cosmid-containing strains tested produced 4 mm colonies within 4 days on medium containing the permissive carbon source fructose.

Construction of mutagenized plasmids

Cosmids were mutagenized with Tn*5* using previously described techniques (Glazebrook and Walker, 1991). In hundreds of separate matings, pLW200 was conjugally transferred into MT614, which carries a Tn*5* insertion in *malE*. Cosmids were then conjugally transferred into MT609, selecting with Sp, Tc and Km to obtain isolates carrying mutagenized cosmids. Individual colonies were picked and used as donors in triparental matings with *A. eutrophus* H16. *A. eutrophus* transconjugants were selected on MM1N containing fructose, Tc and Km, and then challenged to grow on MM1N containing sucrose, Tc and Km. Approximately 10% of the mutagenized cosmids were unable to promote growth on sucrose. Cosmids were conjugally transferred from *A. eutrophus* H16 into *E. coli* C2110 for DNA preparation and analysis. Cosmid-borne Tn*5* insertions were localized using standard restriction mapping techniques.

Sugar transport assays

[U-^{14}C]sucrose (615 mCi mmol^{-1}; 22.8 GBq mmol^{-1}) was obtained from Amersham Life Science (Buckinghamshire, England). *A. eutrophus* H16 harboring pLAFR1, pLW200 or pLW249 was grown in MM1 Tc supplemented with both 0.5% fructose and 0.4% sucrose until the OD_{600} was approximately 0.8. Cells were pelleted and resuspended in MM1 Tc 0.4% sucrose and incubated for 2 hours at 30°C, after which they were pelleted, resuspended in an equal volume of MM1 and incubated without carbon source for 15 min at 30°C. 1 ml was withdrawn for measurement of optical density, and radiolabeled sucrose was added to 2 ml of cells to a final concentration of 1.6 nM. Samples (0.1 ml) were withdrawn in duplicate at 0.5, 3, 5, 7.5, 30 and 60 min after the addition of label, applied to Millipore HA filters under vacuum and washed twice with 2 ml MM1. Filters were dried at 68°C for 15

minutes and 5 ml Hydrofluor scintillant (National Diagnostics, Atlanta) was added before dpm were determined in a Beckman LS 6000SC scintillation counter.

Genetic mapping techniques

The genes identified in this study were mapped using the method of Finan et al. (Finan et al., 1986). Genomic DNA from Rm1021 and *Agrobacterium tumefaciens* strains At123, At125 and At128 was digested with *Eco*RI, subjected to electrophoresis in a 0.6% agarose gel, transferred to a Gene Screen Plus membrane and probed with the insert from pLW201, which contains the C terminal half of *aglG* and the majority of *aglA*.

Plant inoculation assays

Medicago *sativa* cv. Iroquois was obtained from Agway (Plymouth, IN). *R. meliloti* strains were tested for the ability to nodulate alfalfa on nitrogen-free Jensen's medium as described (Leigh et al., 1985). Plants were grown in a constant temperature room at 25°C with a 20 hour light cycle. Observations were made weekly for at least six weeks. Each nodulation assay included the control treatments of water, wild type *R. meliloti* Rm1021 and an *exoA* mutant Rm7031 (Leigh et al., 1985). The presence of pink, cylindrical nodules on dark green healthy plants was taken as evidence that nitrogen fixation was occurring. Plants lacking nodules or with ineffective nodules were stunted and chlorotic.

RESULTS

Identification of R. meliloti *genomic clones which promote growth on sucrose*

A *R. meliloti* genomic library was mated into wild type *A. eutrophus*, and transconjugants were selected on MM1N minimal medium containing abundant nitrogen and sucrose as the sole carbon source. Approximately 24,000 transconjugants were screened from three separate matings. Less that 0.1% of the transconjugants were able to grow on the sucrose plates. We scored colonies for growth by comparing the size of colonies with an isogenic strain harboring the vector. Although *A. eutrophus* cannot utilize sucrose as a carbon source, it does form microcolonies on this medium because it is a facultative chemolithotroph capable of using CO_2 as a carbon source. These translucent microcolonies were only seen when plates were examined with a dissecting microscope.

In contrast, some transconjugant strains produced visible colonies after 3 days of incubation, and we scored for ability to grow on sucrose after 5 days of incubation. These colonies have the same color and colony morphology as strains grown on fructose. No cosmids were identified which allowed *A. eutrophus* to grow at the same rate on sucrose plates as on fructose, and strains carrying cosmids of interest reached 4 mm in diameter only after 11 days of incubation. Seventeen colonies which permit the growth of the recipient on sucrose were selected for further study.

Identification of a segment of R. meliloti *DNA that improves the growth of* A. eutrophus *on sucrose*

Restriction mapping revealed that the seventeen cosmids chosen appear to contain the same region of the *R. meliloti* genome. Only two restriction patterns were represented in the candidate cosmids. A 19.4 kb region of DNA is present in each of the cosmids, and two of the cosmids contain an additional 3 kb of DNA. We chose a

representative cosmid carrying the smaller insert (pLW200) for further study (Figure 4-1A). No single *Eco*RI restriction fragment from pLW200, when cloned into a broad host range vector, was able to confer the growth advantage on sucrose. To identify the region(s) of importance, pLW200 was subjected to transposon Tn*5* mutagenesis and then mated into *A. eutrophus*. Transconjugants were selected on MM1N fructose plates and then tested for their ability to grow on sucrose. Approximately 10% of the transconjugants carrying mutagenized cosmids were unable to grow on sucrose. In all, more than seventy mutagenized cosmids were isolated which failed to promote growth on sucrose. The Tn*5* insertions in these cosmids map to a central 6.8 kb region which is required for the improved growth phenotype (Figure 4-1C).

The same segment of R. meliloti *DNA also permits growth of* A. eutrophus *on maltose and trehalose*

In order to relate our observations to the previous study of disaccharide metabolism by *R. meliloti* (Glenn and Dilworth, 1981), we tested whether the cosmid conferring the ability to grow on sucrose influenced the ability of *A. eutrophus* to grow on maltose, trehalose or the galactosides lactose or melibiose. The cosmid pLW200 improved the growth of *A. eutrophus* on the α-glucoside disaccharides, but not on lactose or melibiose. Nor did it influence growth on fructose. pLW249, one of the Tn*5* insertion mutants of pLW200 identified because it fails to improve the growth of *A. eutrophus* on sucrose, is also unable to promote growth on maltose or trehalose as a sole carbon source. These results are consistent with the suggestion (Glenn and Dilworth, 1981) that *R. meliloti* uses the same transport system for sucrose, trehalose and maltose.

DNA sequence suggests this region encodes a binding-protein dependent transport system for alpha-glucosides

Analysis of transposon insertions and DNA sequence data has led to the detection of 5 loci within the 6.8 kb region identified by Tn*5* mutagenesis (Figure 4-1B). Since they have been implicated in α-glucoside utilization, we have named these genes *agl*. The genes are arranged in the order *aglEFGAK*. On the basis of sequence homologies *aglA* appears to encode an α-glucosidase, whereas *aglK*, *aglF*, *aglG* and *aglE* appear to encode a periplasmic binding protein dependent transport system (discussed below). Transport of the α-glucosides sucrose, maltose and trehalose followed by their cleavage by AglA can account for how this set of genes confers the ability to utilize these three sugars.

Mutants of R. meliloti *disrupted by insertions in genes carried on pLW200 can utilize sucrose and are Fix+*

All of the Tn*5* insertions shown in Figure 4-1C were transferred from the mutagenized cosmid into the Rm1021 genome by homogenotization and the Tn*5* insertions were subsequently transduced into Rm1021 to ensure a clean genetic background. The resulting mutants were tested for their ability to utilize various carbon sources. All of the mutants are able to utilize glucose and succinate as carbon sources. *R. meliloti aglA* mutants, which are disrupted in the gene encoding the putative glycanase, grew as well as or almost as well as wild type in sucrose and trehalose but grow more slowly in liquid cultures containing maltose. *R. meliloti aglE*, *aglF* and *aglG* mutants, which are disrupted in the putative periplasmic binding protein and inner membrane permeases, respectively, grow more slowly than wild type or *aglA* in liquid cultures containing sucrose, maltose or trehalose and fail to grow in liquid cultures in which the concentration of trehalose or

maltose was reduced to 1 mM. *aglE*, *aglF* and *aglG* mutants form weakly growing colonies on M9 plates containing sucrose, maltose or trehalose. These data strongly suggest that *R. meliloti* possesses at least one additional α-glucosidase activity besides the one proposed to be encoded by *aglA*. Because the mutants defective in the putative transport system are more severely affected than those affected in the putative glycanase, we propose that *aglE*, *aglF*, *aglG* and *aglK* encode the major transport system for import of α-glucosides into the cell, and that there is at least one other, possibly lower affinity, pathway for import of these sugars. In addition, because the growth defect is more severe in the *aglE*, *aglF* and *aglG* mutants than the *aglA* mutant, these data suggest that the proposed additional α-glucosidase is not an extracellular enzyme. All of the *agl* mutants are able to elicit Fix^+ nodules on *Medicago sativa*.

AglA shares homology with α-glucosidases

The predicted amino acid sequence of AglA shows homology to many members of family 13 of glycanases, also referred to as the α-amylase family (Henrissat, 1991; Henrissat and Bairoch, 1993; Henrissat and Romeu, 1995). The family is composed of proteins with diverse substrate specificities and products; enzymes which may cleave maltose, sucrose, trehalose and oligosaccharides composed of glucose with either α-1,4 or α-1,6 linkages. The strongest homology is seen with an exo-α-1,4 glucosidase from *Bacillus stearothermophilus* (Takii et al., 1996), a thermostable α-glucosidase of *Bacillus* sp. (Nakao et al., 1994) and an oligo-1,6-glucosidase from *Bacillus coagulans* (Watanabe et al., 1996). Strong homology is also observed with many maltases, sucrases and two trehalose hydrolase proteins. The deduced polypeptide of AglA contains the seven invariant amino acid that are conserved in the active site of enzymes of the α-amylase family (Figure 4-2) (Svensson, 1994).

aglK *appears to encode the ABC component of the binding protein dependent transport system*

The deduced polypeptide AglK is homologous (55% identical/69% similar) to *Agrobacterium radiobacter* LacK (Williams et al., 1992) and shows strong homology to other members of the ATP-binding cassette (ABC) family of proteins. An alignment of AglK with *A. radiobacter* LacK and *E. coli* MalK, the ATP-binding component of the maltose transport system, is shown in Figure 4-3. There are two strongly conserved motifs located within the N-terminal region of the protein, and the degree of similarity drops off sharply in the C-terminus. An ATP/GTP binding site, or P-loop, is located at residues 36-43. The highly conserved ABC transporter signature, a PROSITE pattern found in all proteins of this family, is located at residues 134-144 of the deduced amino acid sequence.

AglF and AglG are homologous to inner membrane sugar permeases

AglF and AglG share homology with the MalFG subfamily of inner membrane permeases and contain the PROSITE pattern developed as a signature of this family of proteins (Figure 4-4 and Figure 4-5). In the homologous *E. coli* protein MalG, the protein domain defined by this consensus sequence has been proposed to be involved in protein-protein interactions with MalK, a cytoplasmic ATP-hydrolyzing peripheral membrane protein and a homologue of AglK (Dassa, 1993; Nikaido, 1994). The strongest homologues of AglF and AglG are Slr0530 and Slr0531 respectively, hypothetical proteins of *Synechocystis* sp. strain PCC6803 that are thought to be inner membrane permeases (Hirosawa et al., 1995; Kaneko et al., 1996). AglF is 52% identical and 73% similar to Slr0530. The hydrophobicity trace of AglF is almost superimposable with those of Slr0530 and *A. radiobacter* inner membrane lactose permease LacF (Williams et al., 1992). There are six potential

transmembrane domains, consistent with the hypothesis that it is an integral membrane protein. AglG is 52% identical and 72% similar to Slr0531, and the hydrophobicity trace of AglG suggests that it could form 5 or 6 transmembrane domains. Taken together it seems likely that AglF and AglG are integral membrane proteins which may be involved in transport of sugar substrates across the inner membrane.

AglE shares homology with periplasmic solute binding proteins

AglE appears to be a member of the MalE periplasmic solute binding protein family. AglE is 42% identical and 62% similar to Slr0529, a hypothetical protein of *Synechocystis* sp. strain PCC6803, which in turn shares homology with periplasmic solute binding proteins MalE of *E. coli* and MsmE of *Streptococcus mutans* (Figure 4-6). Both of these proteins specifically bind sugars; MalE is a maltose/maltodextrin binding protein (Shuman, 1982; Duplay et al., 1984) and MsmE is involved in the uptake of melibiose, raffinose and isomaltotriose (Russell et al., 1992). MalE, the *E. coli* maltose binding protein, has been extensively analyzed. Structure-function and crystallization studies have shown that the substrate binding site lies in a cleft between the N and C terminal lobes of the protein (Spurlino et al., 1991). The regions which are conserved between these known periplasmic solute binding proteins and AglE and Slr0529 are limited. However, the highest conservation between MalE, MsmE, Slr0529 and AglE is in a region which in MalE forms the hinge between N and C terminal domains and is adjacent to residues which have been shown to contact the substrate in the ligand bound crystal (Spurlino et al., 1991). The structure of maltose binding protein has been compared with that of other periplasmic substrate binding proteins, and although they tend to have very

different primary sequences, their 3-dimensional structures share many similarities (Quiocho and Ledvina, 1996).

Support for this inference that AglE is a periplasmic binding protein is provided by gene order in the *agl* region. In almost every reported case (Boos and Lucht, 1996), genes encoding periplasmic solute binding proteins are directly upstream of their associated inner membrane permease genes. One example is the lactose operon of *A. radiobacter* (Williams et al., 1992) which encodes a binding-protein dependent periplasmic lactose permease and most closely resembles the gene order of the *agl* region. The gene encoding lactose binding protein, *lacE*, is found upstream of *lacF* and *lacG* encoding integral membrane sugar permeases, *lacZ* encoding β-galactosidase and *lacK* which encodes the ATP-hydrolyzing component of the transporter. A divergently transcribed regulatory locus is located upstream of *lacE*. The gene order of the *agl* region duplicates this gene order, and it seems possible that *aglE* could encode a periplasmic binding protein which would interact with the putative sugar permeases encoded by *aglF* and *aglG*.

AglR, a putative regulatory protein, is homologous to DNA binding proteins

Upstream of *aglE* and divergently transcribed is a locus which shows homology to transcriptional regulators of the *lacI* family of repressors and which was named *aglR* on the basis of this homology. Many proteins in this family are involved in catabolite repression of sugar utilization operons. The strongest homologue of the deduced polypeptide AglR is RafR, the raffinose regulator of *E. coli* (Aslanidis et al., 1989). AglR is 32% identical and 50% similar to RafR, and an alignment of these deduced polypeptides with the regulatory protein CcpA of *Bacillus megaterium*, which negatively regulates expression of an amylase and positively regulates genes involved in the excretion of excess carbon (Hueck et al., 1994), and the negative

transcriptional regulator CytR of *E. coli* (Valentin-Hansen et al., 1986) is shown in Figure 4-7. *aglR* lies outside of the 6.8 kb region of pLW200 shown to be required for the utilization of α-glucoside disaccharides by *A. eutrophus*, as would be expected if it serves as a negative regulator of the expression of the other *agl* genes.

The agl *genes permit uptake of sucrose by* A. eutrophus

To test our model that the *aglE*, *aglF*, *aglG* and *aglK* gene products are involved in transport of α-glucosides, we examined the ability of an *A. eutrophus* strain harboring pLW200 to incorporate radiolabeled sucrose. These experiments were conducted in *A. eutrophus* in order to observe the activity of the *agl* genes outside of the context of other *R. meliloti* sucrose transport or hydrolysis systems. As shown in Figure 4-8, pLW200, which carries the *agl* genes, is able to promote the uptake of ^{14}C-sucrose. This effect is not seen in an isogenic strain carrying the vector pLAFR1, nor in a strain harboring pLW249, a derivative of pLW200 which has a Tn*5* insertion in the *aglE* gene. These results suggest that the genes carried on pLW200 encode a functional system for transport of sucrose.

The agl *region maps to the* R. meliloti *chromosome*

R. meliloti has three replicons, the chromosome and two megaplasmids of 1.4 and 1.7 Mb (Sobral et al., 1991). The *agl* region was mapped by Southern hybridization using the method of Finan et al. (Finan et al., 1986). *Eco*RI-digested genomic DNA from Rm1021, Rm9623, Rm9624, Rm9625 and *Agrobacterium tumefaciens* strains cured of the Ti plasmid and carrying (a) *R. meliloti* megaplasmid pRmeSU47a, (b) *R. meliloti* megaplasmid pRmeSU47b or (c) no megaplasmid was probed with radiolabeled pLW201 DNA, which contains the C-terminal half of *aglG* and the majority of *aglA*. A strongly hybridizing band of 2.1 kb was observed in the lane

containing DNA isolated from Rm1021, and strongly hybridizing bands of approximately 8 kb were seen in the lanes containing DNA from the three Tn*5* insertion mutants. No bands were seen in the lanes containing DNA from the *A. tumefaciens* strains, indicating that the locus maps to the *R. meliloti* chromosome and not to one of the megaplasmids.

zwf *and* edd

Because all of the cosmids identified in our screen had such a large (~20 kb) overlap we were interested in whether additional loci involved in sugar metabolism mapped nearby, and we determined the DNA sequence of 7.3 kb downstream of the *agl* region. Several open reading frames were identified. Of particular interest are two loci which appear to encode the first two enzymes of the Entner-Doudoroff pathway (Figure 4-9). One open reading frame encodes a deduced protein of 491 amino acids that is 51% identical and 67% similar to *Zymomonas mobilis* glucose-6-phosphate dehydrogenase enzyme Zwf (zwischenferment) (Barnell et al., 1990) and contains the PROSITE pattern DHYLGK[E/Q] which is conserved among glucose-6-phosphate dehydrogenase proteins. Based on this information, we have named this locus *zwf*. An alignment of *R. meliloti* Zwf and the Zwf proteins from *Z. mobilis*, *E. coli* and *Synechocystis* sp. is shown in Figure 4-10. An second open reading frame encodes a deduced protein of 605 amino acids whose strongest homologue (58% identity/75% similarity) is *A. tumefaciens* MocB (Kim and Farrand, 1996), a homologue of 6-phosphogluconate dehydratase (Edd), the Entner Doudoroff enzyme which converts 6-phosphogluconate to 2-keto-3-deoxy-6-phosphogluconate. The *R. meliloti* deduced protein sequence contains two PROSITE signatures (Bairoch, 1992) which are conserved in the evolutionarily related dihydroxy acid dehydratases and 6-phosphogluconate dehydratases (Figure

4-11) and we have named the locus *edd*. The Entner-Doudoroff pathway is known to be the major pathway for glucose utilization in *R. meliloti* (Stowers, 1985), but these loci have not been previously cloned from *R. meliloti*. The activities encoded by *zwf* and *edd* have been detected in the bacteroid fraction of alfalfa nodules (Irigoyen et al., 1990). A nitrosoguanidine-induced mutant of *R. meliloti* lacking glucose-6-phosphate dehydrogenase activity has been isolated (Cerveñansky and Arias, 1984) and was found to be Fix^+ on alfalfa, suggesting that this enzymatic activity is not essential during symbiosis or that there may be an additional, developmentally regulated locus encoding this activity.

R. meliloti *DevB is a member of the SOL/DevB family of oxidoreductases*

A small open reading frame between *zwf* and *edd* encodes a polypeptide of 188 amino acids with homology (33% identity/51% similarity) to DevB from the cyanobacterium *Anabaena* sp. strain 7120 (Figure 4-12), and we have provisionally named this locus *devB*. Although *Anabaena devB* has been proposed to encode an isozyme of glucose-6-phosphate dehydrogenase (Bauer, 1994) based on its homology with the central portion of a biochemically characterized isozyme of glucose-6-phosphate dehydrogenase isolated from rabbit liver microsomes (Ozols, 1993), this designation seems premature, because the region of homology between *Anabaena* DevB and the rabbit glucose-6-phosphate dehydrogenase does not include the catalytic domain of the protein. The conserved active site sequence DHYLGK[E/Q] found in all homologues of Zwf is absent in DevB. Instead Shen et al. offer the model that the rabbit microsomal glucose-6-phosphate dehydrogenase is a chimeric protein composed of a Zwf domain and a SOL domain and define a new family of proteins, the SOL/DevB family of putative oxidoreductases (Shen et al., 1996). We propose that the *R. meliloti* deduced polypeptide DevB is a member of this family of

proteins. *Anabaena devB* was identified in a subtractive hybridization screen and is preferentially expressed in heterocysts, the differentiated cells of *Anabaena* where nitrogen fixation takes place (Bauer, 1994). It would be interesting to determine whether *R. meliloti devB* is differentially regulated in the nitrogen-fixing bacteroid form.

The 5' terminus of the DNA sequence presented in this chapter encodes a truncated open reading frame immediately upstream of *zwf*. This partial open reading frame encodes 305 amino acids which are 27% identical and 50% similar to the C-terminal portion of *Haemophilus influenzae* probable oxidoreductase OrdL, and we have provisionally named this locus *ordL* (Figure 4-13).

Repeat elements and an insertion element in the agl *region*

An unusual feature of this region is the presence of two copies of RIME1, a *Rhizobium*-specific intergenic mosaic element. This repeat element was first identified between *chvI* and *exoS* of *R. meliloti* (Østerås et al., 1995). RIME1 contains two large inverted repeats, and was named for its structural similarity to BIME, a repeat element found in enteric bacteria (Gilson et al., 1991). The 5' copy of RIME1 is found between *aglE* and *aglF*, and the 3' copy is immediately upstream of the insertion sequence described in the next section.

RIMEs have been previously identified by DNA sequencing in *R. meliloti*, *Rhizobium* spp. NGR234 and *Rhizobium leguminosarum* (Østerås et al., 1995). In addition, RIME sequences hybridized to DNA from *Agrobacterium rhizogenes*, but not that of *A. tumefaciens* or *Bradyrhizobium japonicum* (Østerås et al., 1995). Although Østerås et al. reported than no copies of RIME were identified on the *Rhizobium* spp. NGR234 megaplasmid by Southern hybridization, we were able to identify three copies of RIME1 by performing a BLASTN search of the symbiotic

megaplasmid of *Rhizobium* NGR234 (Freiberg et al., 1997). We also identified additional copies in *Rhizobium trifolii* and the *phoCDET* (Bardin et al., 1996) and *exp* (Becker et al., 1997) regions of *R. meliloti*. All copies of RIME1 found to date are located in intergenic regions or overlapping the coding region by a few bases.

We have also identified a novel insertion element between *edd* and *aglK*. This insertion element consists of an apparently fragmented reading frame flanked by 39 bp terminal inverted repeats. The nucleotide sequence in this region is highly homologous to the *Rhizobium* spp. NGR234 megaplasmid locus *y4zb*, and the *R. meliloti* potential coding region encodes a polypeptide with strong homology (interrupted by in-frame stop codons) to the hypothetical protein Y4zb. Y4zb is thought to be a transposase, but the *y4zb* locus is not flanked by the inverted repeats seen in *R. meliloti*. Both Y4zb and its *R. meliloti* homologue exhibit limited homology to transposases from *Bacillus stearothermophilus* (Xu et al., 1993) and insertion sequences and we have therefore tentatively named the locus *tnp*. It is possible that this locus may be, or may have been involved in integration or recombination functions at one time during the evolution of the strain.

The 39 bp terminal sequences show homology with intergenic sequences in *R. meliloti* and *R. trifolii*. We observed that this 39 bp sequence is 100% conserved in an intergenic region downstream of *R. meliloti ftsZ* (Margolin and Long, 1994). We also found that the same 39 bp are 87% conserved with a sequence between the *R. meliloti* betaine aldehyde dehydrogenase (*betB*) and choline dehydrogenase (*betA*) genes (Pocard et al., 1997), and 86% conserved with sequence overlapping the stop site of the *Rhizobium trifolii* fructokinase (*frk*) gene (Fennington and Hughes, 1996). These data are shown in Figure 4-14. It seems possible that these occurrences of the 39 bp repeat element represent former sites of genomic recombination.

DISCUSSION

We have isolated a cosmid containing *R. meliloti* DNA which is able to promote growth on sucrose, maltose or trehalose and have constructed Tn*5* insertions in the cosmid which abolish that ability. DNA sequencing led to the identification of five possible protein coding regions within the boundaries defined by Tn*5* insertions. We have demonstrated that the cosmid promotes the uptake of radiolabeled sucrose by *A. eutrophus*, and that *R. meliloti agl* mutants are affected for growth on α-glucosides. Considering these data and incorporating inferences made about these proteins on the basis of their deduced amino acid sequences, we propose that the AglEFGK proteins form a periplasmic binding protein dependent transport system for α-glucoside disaccharides and that AglA is a glycosyl hydrolase active on the disaccharides trehalose, sucrose and maltose. Although *agl* mutants grow more slowly on α-glucoside disaccharides, they can still utilize these carbon sources and our model therefore accounts for these observations by including at least one additional mechanism for transporting and cleaving sucrose, maltose and trehalose (Figure 4-15.).

If our model is correct and *R. meliloti* possesses more than one glycanase which is able to cleave sucrose, maltose and trehalose, it would not be the first example of redundant glycanase activity in *R. meliloti*. *exoK* mutants, which are deficient in the production of an extracellular glycanase active on the acidic exopolysaccharide succinoglycan, still possess succinoglycan-cleaving activity. Three loci, *exsH*, *prsD* and *prsE*, were found to encode the second glycanase and the system required to export the glycanase (York and Walker, 1997). The *R. meliloti* genome is approximately 6.8 Mb, several fold larger than the genome of *Haemophilus influenzae* (Fleischmann et al., 1995), so the existence of redundant functions is perhaps not too surprising.

Periplasmic binding protein dependent transport systems

The ABC (ATP-binding cassette) transporter superfamily contains many exporters (Fath and Kolter, 1993) such as the multi-drug resistance pumps of eukaryotes and the *E. coli* alpha hemolysin exporter. A large evolutionarily related (Saurin and Dassa, 1994) subfamily of ABC transporters are the bacterial periplasmic permeases, also called periplasmic binding-protein dependent transport systems (Boos and Lucht, 1996). These systems carry out energy-dependent transport of substrates across the inner membrane of bacterial cells. Known substrates for bacterial periplasmic permeases include sugars, amino acids, peptides, vitamins and metals (Boos and Lucht, 1996).

The canonical member of the periplasmic binding protein dependent transport systems is the maltose/maltodextrin transporter of *E. coli* (Nikaido, 1994). Bacterial permease systems all contain a hydrophilic peripheral membrane protein which contains the ATP binding cassette, such as *E. coli* MalK (Bavoil et al., 1980). This ABC component is characteristically associated with two hydrophobic integral membrane permeases such as MalF and MalG. The final component is a periplasmic substrate binding protein, which recognizes the substrate and delivers it to the membrane. Proteins in this final class, typified by the maltose binding protein MalE, are comprised of two lobes, separated by a flexible hinge, and a deep binding cleft. The lobes make independent contacts with the two different inner membrane permeases (Hor and Shuman, 1993; Szmelcman et al., 1997) and the cleft is the site of substrate binding (Scharff et al., 1992).

Periplasmic binding protein dependent systems in R. meliloti

The *agl* genes encode one of the first periplasmic binding protein dependent transport systems identified in *R. meliloti*. Gage and Long have only recently

identified genes encoding a *R. meliloti* periplasmic binding protein dependent transport system for α-galactosides (Gage and Long, Personal communication). The *osp* (oligosaccharide permease) genes are homologous to *opp* genes encoding oligopeptide permeases (Hiles et al., 1987). Tn*phoA* insertions in the *osp* genes render *R. meliloti* unable to grow on α-galactosides but do not affect the symbiotic properties of the strain. To test whether removal of both the *osp* and *agl* systems would affect symbiosis, we constructed double mutants which are disrupted in the genes encoding the α-galactoside binding protein (*ospA*) and either the proposed α-glucoside binding protein (*aglE*) or glycosyl hydrolase (*aglA*). Both the *ospA aglE* and *ospA aglA* strains are still able to utilize α-glucosides and are Fix$^+$. In addition to the periplasmic binding protein dependent systems for α-glucosides and α-galactosides, *R. meliloti* appears to contains a periplasmic binding protein dependent system for fructose uptake. Williams et al. showed that an *R. meliloti* periplasmic protein cross-reacts with an antibody raised against the *A. radiobacter* fructose binding protein (Williams et al., 1995). However, the genes encoding the presumed fructose transport system have not been identified.

One of the most novel inferences to come out of this work is that *aglE* may encode a periplasmic binding protein with specificity for sucrose, maltose, and trehalose. If our model is correct, to our knowledge we have identified the first bacterial periplasmic binding that specifically binds to sucrose, maltose and trehalose. Although the deduced protein AglE shares only weak homology with *E. coli* MalE, the region of highest homology is adjacent to residues known to be involved in substrate binding. Our model also implies that Synechocystis proteins Slr0529, Slr0530 and Slr0531 are involved in sugar transport. Sucrose and trehalose are known to be transported by *Synechocystis* (Mikkat et al., 1997), but no genes have yet been assigned to those functions.

The high affinity binding of MalE (MBP) to amylose resin is well documented and this property has been frequently employed for the purification of fusion proteins in overexpression systems (Kellerman and Ferenci, 1982; Bedouelle and Duplay, 1988). One way to demonstrate that AglE is a periplasmic α-glucoside binding protein would be to isolate periplasmic proteins from cells grown in minimal medium containing maltose and determine which proteins bind to an amylose resin. Preliminary results using periplasmic proteins isolated from Rm1021 show that two prominent proteins are retained by the amylose resin and elute with 10 mM maltose. The critical experiment will be to repeat the procedure with an *aglE* mutant strain to determine whether one of the maltose binding proteins is absent. The presence of more than one periplasmic maltose binding protein is consistent with, but not a requirement of, our model that *R. meliloti* has more than one system for the utilization of α-glucosides. The region of starch-filled cells in alfalfa nodules raises the possibility that the starch breakdown product maltose could be a carbon source for bacteroids.

Trehalose is a well characterized osmolyte, and is known to accumulate in highly stressed cultures of *R. meliloti* in which the concentration of NaCl is above 0.5 M (Smith et al., 1994). le Rudulier has recently identified a *R. meliloti* strain unable to utilize trehalose but which is still able to transport trehalose (le Rudulier, personal communication). This strain may carry a mutation in one of the uncharacterized systems for utilization of α-glucosides by *R. meliloti*. It will be important to construct a strain carrying the mutation described by le Rudulier and an *agl* mutation in order to determine whether the double mutant can utilize trehalose.

Our results show that the cosmid-borne *agl* genes are expressed in *A. eutrophus* and that *R. meliloti* strains carrying mutations in *agE* and *aglF* are affected in their

growth on maltose and trehalose. These data show that *agl* genes are involved in *R. meliloti* α-glucoside metabolism. However, we have not examined the transcriptional regulation of the *agl* genes. It would be informative to construct transcriptional *lacZ* fusions to *agl* structural genes and assay for β-galactosidase activity when cells are grown in the presence or absence of substrate. These experiments could also test whether the *agl* genes are subject to catabolite repression by examining β-galactosidase activity in the presence of α-glucosides with or without added succinate, which has been shown to be involved in catabolite repression in *R. meliloti* (Ucker and Signer, 1978). In addition, a double mutant *aglR aglA::lacZ* should be constructed to test the hypothesis that *aglR* is a transcriptional regulator. The fact that none of the Tn*5* insertions which knock out the ability of pLW200 to promote growth on alpha glucosides were located in *aglR* is consistent with the hypothesis that *aglR* encodes a repressor, because elimination of a repressor could lead to constitutive activation of the *agl* genes. In order to relate our observations to the previous study which showed that sucrose uptake by *R. meliloti* is inducible (Glenn and Dilworth, 1981), it would also be informative to examine uptake of radiolabeled sucrose in *R. meliloti* strains carrying insertions in the *agl* genes.

We have characterized the DNA region which lies downstream of the *agl* genes and found several open reading frames which appear to encode enzymes involved in oxidation/reduction reactions. Although only the *agl* genes, located in the central 7 kb of the approximately 20 kb found in all the cosmids isolated in this work, were shown to be required for α-glucoside utilization, the *zwf* and *edd* genes are proposed to be involved in metabolism of the monosaccharides produced after cleavage of sucrose, trehalose and maltose. *zwf* and *edd* are the first examples of

Entner-Doudoroff genes cloned from *Rhizobium, Sinorhizobium, Bradyrhizobium* or *Azorhizobium*. The substrates for the products of *devB* and *ordL* are unknown.

The presence of two copies of RIME1 in this region is intriguing. No one knows the function or origin of RIME1, but it could represent leftover termini from transposition or cross-species lateral transfer events. The inverted repeats of RIME could act as a binding site for a regulatory protein. If RIME1 is able to form hairpins, and these hairpins act as transcriptional terminators, it is possible that premature termination could occur at RIME1, leading to an abundance of truncated transcript containing only *aglE*, and a paucity of full length transcripts. This hypothesis is interesting because it has been shown that *E. coli* periplasmic substrate-binding proteins such as MalE are approximately 30 times as numerous in the cell as their cognate integral membrane permeases (Boos and Lucht, 1996). This disparity is thought to reflect transcriptional differences, and a REP repeat element found between *malE* and *malF* is believed to be involved. Alternatively, the difference could be accounted for by differences in ribosome binding site strength. As more genomes are sequenced, several repeat elements and mosaic elements will undoubtedly be found and may prove useful as tools for taxonomic studies.

Because *R. meliloti* mutants disrupted in *agl* genes encoding the putative transport system are severely affected in their growth on α-glucosides, it appears that *aglEFGK* encode the primary system for transport of these sugars. These results are consistent with the report that *R. meliloti* uses the same uptake system to transport sucrose, maltose and trehalose (Glenn and Dilworth, 1981). Identification of the sucrose utilization system encoded by the *agl* genes should help to make it possible to screen for mutants defective in sucrose utilization and thus evaluate the importance of sucrose metabolism during the nodulation process. Indeed,

preliminary results indicate that it is possible to isolate mutants of either *aglA* or *aglE* that are unable to metabolize the α-glucosides sucrose, maltose and trehalose.

Table 4-1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristics	Source or reference
Strain		
Rhizobium meliloti		
Rm1021	SU47 *str-21*	(Meade et al., 1982)
SG1001	Rm1021 *ospA*::Tn*phoA*	D. Gage
Rm9620	Rm1021 *aglK2*::Tn*5*	this work
Rm9621	Rm1021 *aglE49*::Tn*5*	this work
Rm9622	Rm1021 *aglF95*::Tn*5*	this work
Rm9623	Rm1021 *aglA112*::Tn*5*	this work
Rm9624	Rm1021 *aglG127*::Tn*5*	this work
Rm9625	Rm1021 *aglA279*::Tn*5*	this work
Rm9626	Rm1021 *aglA115*::Tn*5*	this work
Rm9627	Rm1021 *aglA182*::Tn*5*	this work
Rm9628	Rm1021 *aglE192*::Tn*5*	this work
Rm9631	Rm1021 *aglA112*::Tn*5-233*[a]	this work
Rm9632	Rm1021 *aglE49*::Tn*5-233*[a]	this work
Rm9633	Rm9631 *ospA*::Tn*phoA*	this work
Rm9634	Rm9632 *ospA*::Tn*phoA*	this work
Alcaligenes eutrophus		
H16	wild type, Sm^S	(Peoples and Sinskey, 1989b)
Agrobacterium tumefaciens		
At123	GMI9023 ≡ GMI9050 cured of pAtC58 Sm^R, Rifampicin^R	(Rosenberg and Huguet, 1984)
At125	GMI9023 pRmeSU47bΩ5007::Tn*5-oriT* Nm^R/Km^R	(Finan et al., 1986)
At128	GMI9023 pRmeSU47aΩ30::Tn*5-11* Gm^R Sp^R/Km^R Sm^R	(Finan et al., 1986)
Escherichia coli		
C2110	*polA*, Nal^R	B. Staskawicz
DH5α	*supE44 ΔlacU169(ϕ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Clontech
MT614	MM294 *malE*::Tn*5*	T. M. Finan
MT609	*polA1, thy*, Sp^R	T. M. Finan
Plasmid		
pRK600	pRK2013 *npt*::Tn*9*, Cm^R	(Finan et al., 1986)
pPH1JI	IncP, Gm^R, Sp^R	(Beringer et al., 1978)
pR751	IncP, Tp^R	(Meyer and Shapiro, 1980)
pLAFR1	Tc^R, IncP broad host range cosmid vector	(Friedman et al., 1982)
pSW213	Tc^R, IncP broad host range vector	(Chen and Winans, 1991)
pAB2001	Source of promoterless *lacZ*Gm interposon	(Becker et al., 1993)
pUTminiTn*5*Km	Source of Nm^R/Km^R cassette	(de Lorenzo et al., 1990)
pLW200	pLAFR1 derivative carrying *R. meliloti agl* region	this work
pLW249	pLW200 *aglE49*::Tn*5*	this work

[a] These mutants were obtained by replacement of the corresponding Tn*5* insertion with Tn*5-233* (De Vos et al., 1986).

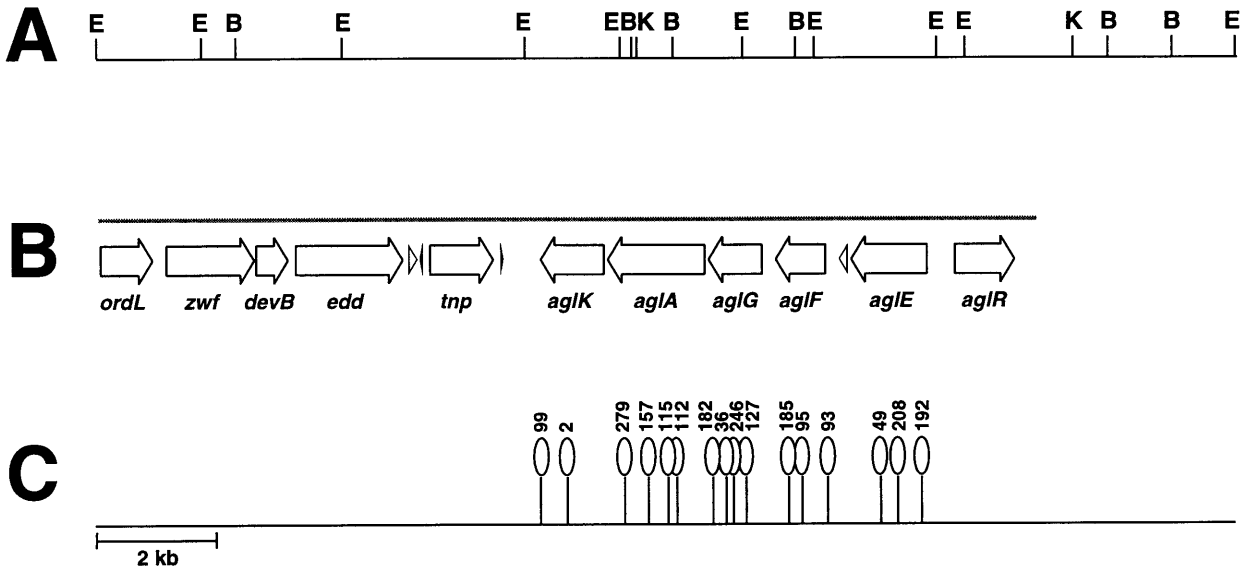


Figure 4-1.

Genetic organization of the *agl* region of *R. meliloti*

A. Restriction map of the insert of pLW200. Restriction sites indicated: *Eco*RI (E), *Bam*HI (B), *Kpn*I (K).

B. Positions and extents of the putative protein coding regions identified by sequence analysis are shown by open arrows. Open triangles indicate the positions and orientation of RIME1. Black triangles indicate the positions of the 39 bp repeat element. The grey line above the arrows indicates the region sequenced.

C. Positions of Tn*5* insertions which eliminate the ability of pLW200 to promote the growth of *A. eutrophus* on the α-glucosides sucrose, maltose and trehalose.


1 MTE---------------WWKKAVVYQIYPRSFYDTNG B. coagulans
1 M----------------KKTWWKEGVAYQIYPRSFMDANG B. stearo.
1 LSTALTQTSTNSQQSPIRRAWWKEAVVYQIYPRSFMDSNG Bacillus sp.
1 MT--MNETTSSLLEA--DRDWWRRPVIYQIYPRSFQDTNG AglA
24 DGIGDLRGIMDKLDYLKTLGIDCIWISPVYDSPQDDNGYD B. coagulans
25 DGIGDLRGIIEKLDYLVELGVDIVWICPIYRSPNADNGYD B. stearo.
41 DGIGDLRGILSKLDYLKLLGVDVLWLNPIYDSPNDDMGYD Bacillus sp.
37 DGIGDLQGITARLPHIAGLGADAIWISPFFTSPMRDFGYD AglA
64 IRDYRKIDKMFGTNEDMDRLLDEAHARGIKIVMDLVVNHT B. coagulans
65 ISDYYAIMDEFGTMDDFDELLAQAHRRGLKILDLVINHT B. stearo.
81 IRDYYKIMEEFGTMEDFEELLREVHARGMKLVMDLVANHT Bacillus sp.
77 VSNYVDVDPIFGTLEDFDALIAEAHRLGLRVMIDLVLSHT AglA
104 SDEHAWFVESRKSKDNPYRDFYFWKDPKPDGTPPNNWGSM B. coagulans
105 SDEHPWFIESRSSRDNPKRDWYIWRDGK-DGREPNNWESI B. stearo.
121 SDEHPWFIESRSSRDNPYRDWYIWRDPK-DGREPNNWLSY Bacillus sp.
117 SDRHPWFVESRSSRSNAKADWYVWADSKPDGTPPNNWLSI AglA
144 FSGSAWEYDETTGQYYLHYFSKKQPDLNWENEKVRKEIYD B. coagulans
144 FGGSAWQYDERTGQYYLHLFDVKQPDLNWENSEVRQALYD B. stearo.
160 FSGSAWEYDERTGQYYLHLFSRRQPDLNWENPKVREAIFE Bacillus sp.
157 FGGSAWQWDPTRLQYYLHNFLTSQPDLNLHNPQVQEALLA AglA
184 MMKFWMDKGVDGWRMDVIGSISKFLDFPDYEL--------- B. coagulans
184 MINWWLDKGIDGFRIDAISHIKKKPGLPDLPN--------- B. stearo.
200 MMRFWLDKGIDGFRMDVINAIAKAEGLPDAPA--------- Bacillus sp.
197 VERFWLERGVDGFRLDTINFYFHDRELRDNPALVPERRNA AglA
216 ---PEGQKYGIG-K-YHANGPRLHAFIQEMNREVLSKY-D B. coagulans
216 ---PKGLKYVPSFA-AHMNQPGIMEYLRELKEQTFARY-D B. stearo.
232 ---RPGERYAWGGQ-YFLNQPKVHEYLREMYDKVLSHY-D Bacillus sp.
237 STAPAVNPYNYQEHIYDKNRPENLEFLKRF-RAVMDEFPA AglA
250 CMTVGEAIGS--DVEIARKYTGPDRHELNMIFNFEHMDVD B. coagulans
251 IMTVGEANGV--TVDEAEQWVGEENGVFHMIFQFEHLGL- B. stearo.
267 IMTVGETGGV--TTKDALLFAGEDRRELNMVFQFEHMDID Bacillus sp.
276 IAAVGEVGDSQRGLEIAGEYTSGG-DKVHMCYAFEFLAPD AglA
288 TKPGSPAGKWALK---PFDLVELKQILSRWQYELADTGWN B. coagulans
288 ----------WKRKADGSIDVRRLKRTLTKWQKGLENRGWN B. stearo.
305 ATDG---DKWRPR---PWRLTELKTIMTRWQNDLYGKAWN Bacillus sp.
315 ----------------RLTPQRVAEVLRDFHRA-APEGWA AglA
325 ALYFENHDQARVVSRWGNDTTYRAECAKAFATILHGLKGT B. coagulans
319 ALFLENHDLPRSVSTWGNDREYWAESAKALGALYFFMQGT B. stearo.
339 SLYWTNHDQPRAVSRFGNDGPYRVESAKMLATVLHMMQGT Bacillus sp.
338 CWAFSNHDVVRHVSRWADGVTDHDAHAKLLASLLMSLRGT AglA
365 PFIYQGEEIGMVNADLE-LEEYDDIEIRNAYQELVMENQI B. coagulans
359 PFIYQGQEIGMTNVQFSDIRDYRDVAALRLYELE-RANG- B. stearo.
379 PYIYQGEEIGMTNCPFDSIDEYRDVEIHNLWRHRVMEGG- Bacillus sp.
378 VCIYQGEELALAEAEL----DYEDLQ-----DPMASSSGP AglA
404 MSKDEFLTAVRKKGRDNARTPMQWDGSFNAGFTTGTPWLK B. coagulans
397 RTHEEVMKIIWKTGRDNSRTPMQWSDAPNAGFTTGTPWIK B. stearo.
418 QDPAEVLRVIQLKGRDNARTPMQWDDSPNAGFTTGTPWIK Bacillus sp.
409 TSSA---------GR--CRTPMVWESLPDGGFSSATPWLP AglA

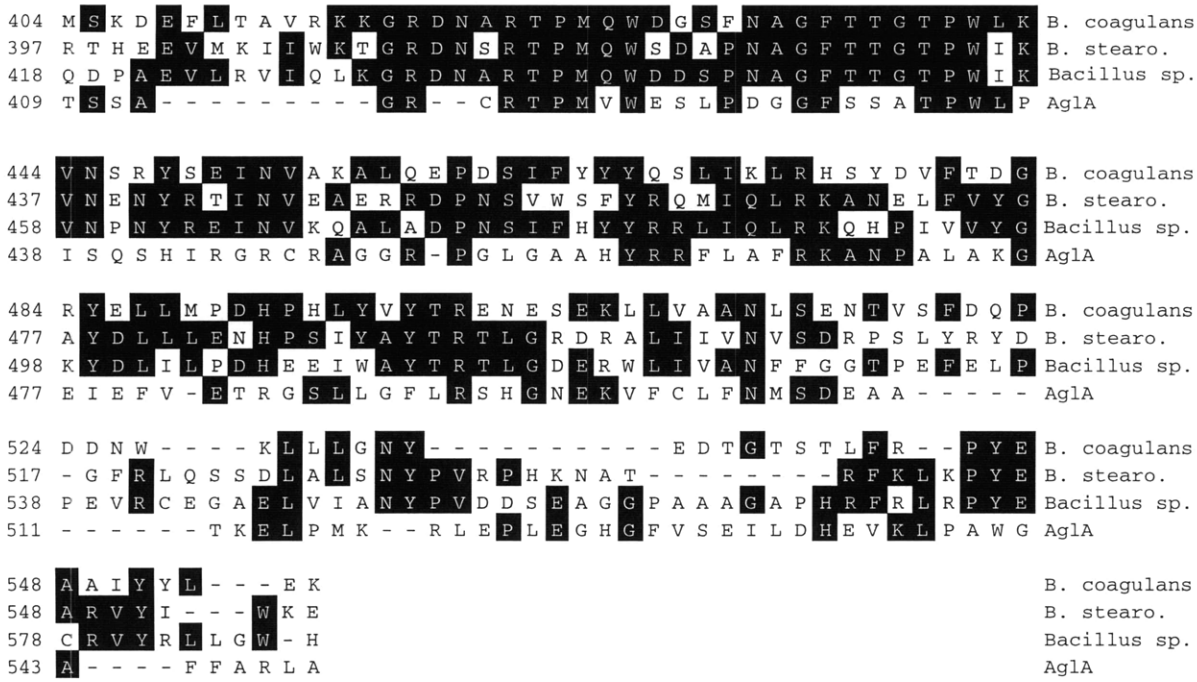


Figure 4-2.
Alignment of the deduced proteins AglA from *R. meliloti*, an exo-α-1,4-glucosidase from *B. stearothermophilus* (1321625), an α-1,4-glucosidase from *Bacillus* sp. (580824) and an oligo-1,6-glucosidase from *B. coagulans* (2492896). Accession numbers are indicated in parentheses. The alignment was performed using the Clustal algorithm of SeqMan (Lasergene). Conserved residues are shaded with black.

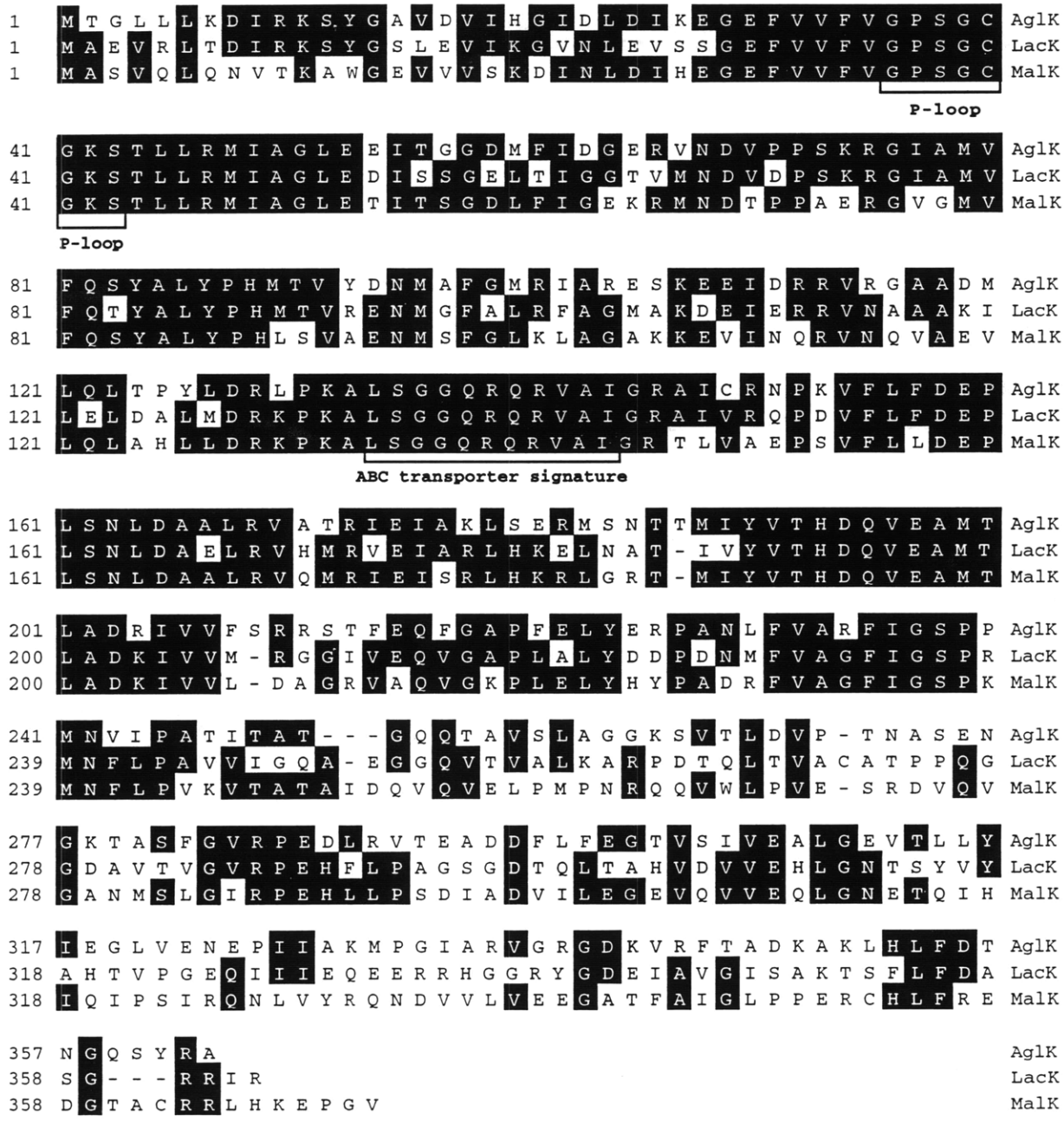


Figure 4-3.
Alignment of the deduced polypeptides AglK of *R. meliloti*, LacK of *A. radiobacter* (266444), and MalK of *E. coli* (409797). The alignment was performed using the Clustal algorithm of Megalign (Lasergene). Gaps have been introduced to optimize the alignment. Identical residues are shaded with black. The PROSITE pattern conserved in ATP-binding cassette peripheral membrane proteins is boxed. Accession numbers are indicated in parentheses.

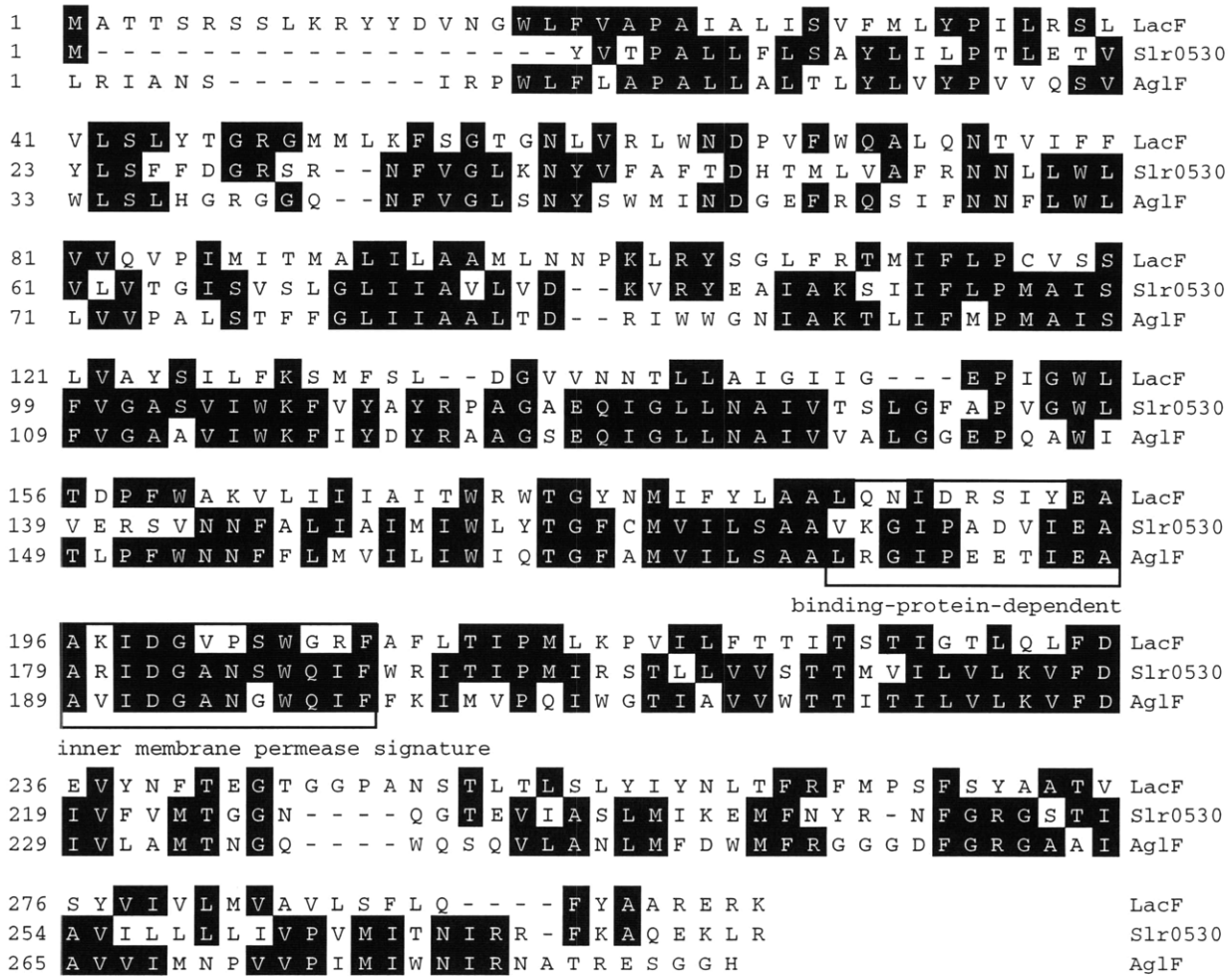


Figure 4-4. AglF

Alignment of the deduced polypeptides AglF of *R. meliloti,* Slr0530 of Synechocystis sp. strain PCC6803 (1001331) and lactose permease LacF of *A. radiobacter* (266441) The alignment was performed using the Clustal algorithm of Megalign (Lasergene). Gaps have been introduced to optimize the alignment. Identical residues are shaded with black. The PROSITE pattern conserved in inner membrane permeases is boxed. Accession numbers are indicated in parentheses.

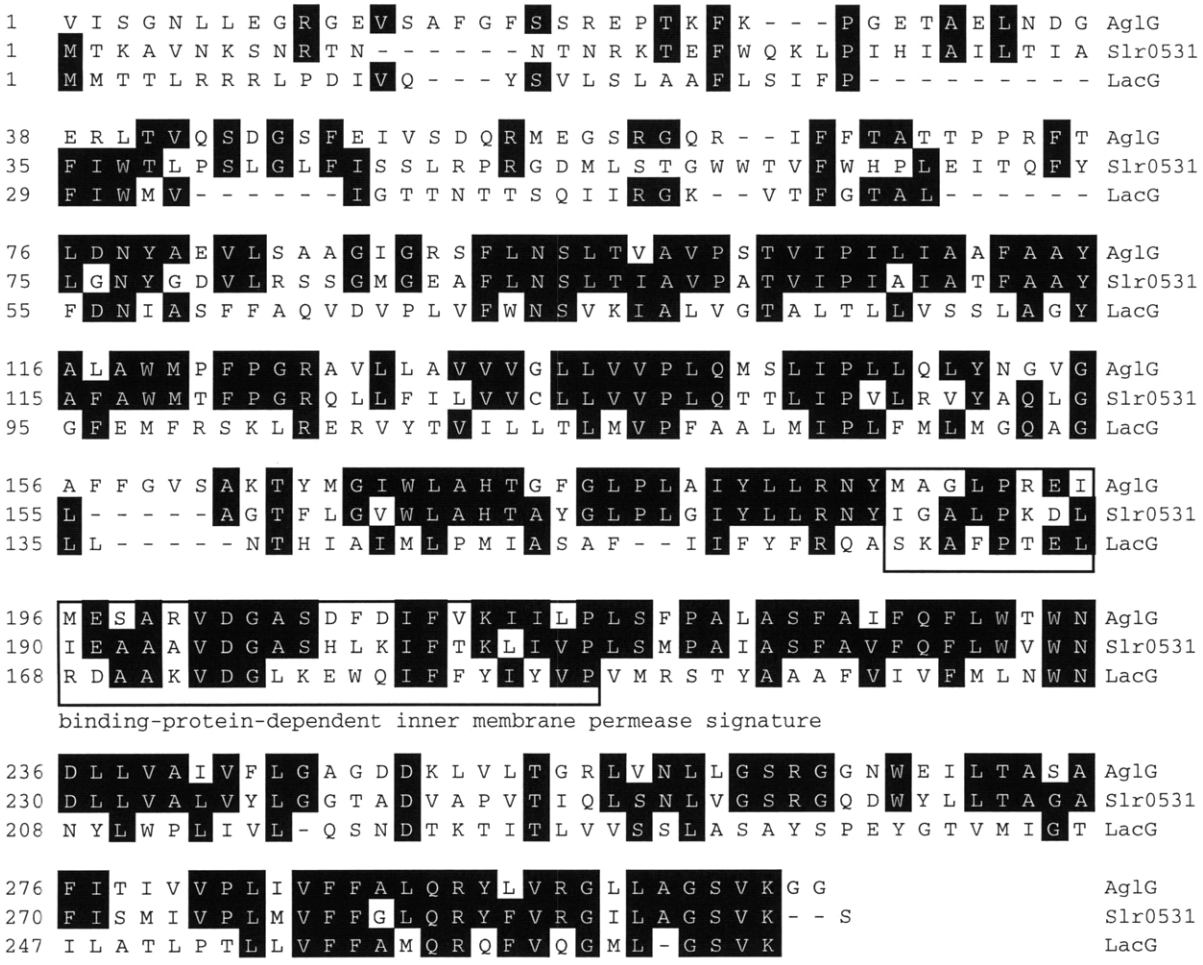
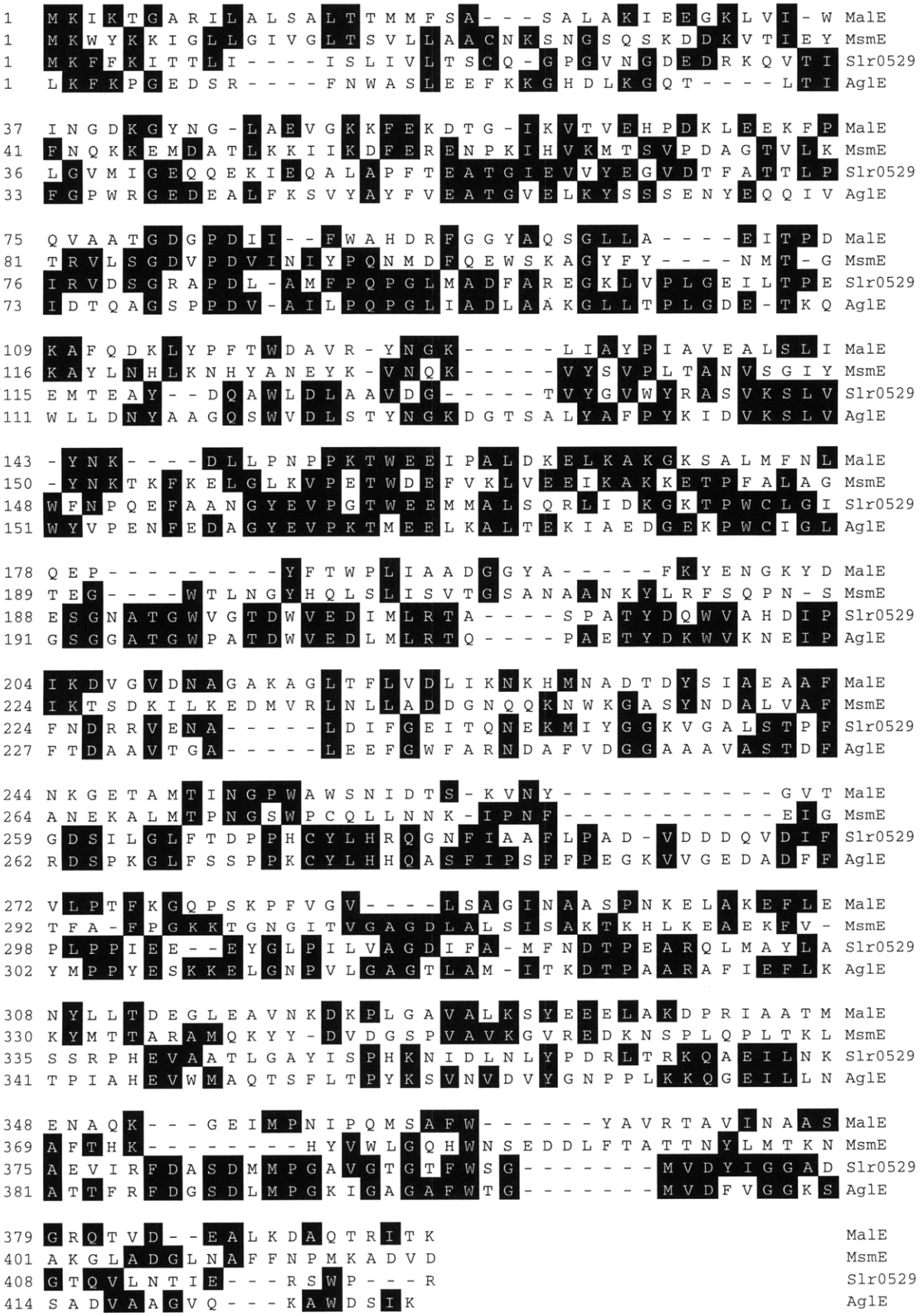


Figure 4-5.
Alignment of the deduced proteins AglG of *R. meliloti,* Slr0531 of *Synechocystis* sp. strain PCC6803 (1001332) and lactose permease LacG of *A. radiobacter* (400157). The alignment was performed using the Clustal algorithm of Megalign (Lasergene). Gaps have been introduced to optimize the alignment. Identical residues are shaded with black. The PROSITE pattern conserved in inner membrane permeases is boxed. Accession numbers are indicated in parentheses.

Figure 4-6.
AglE and Slr0529 are weakly homologous with known periplasmic sugar binding proteins. An alignment of the deduced peptides AglE of *R. meliloti*, Slr0529 of *Synechocystis* sp. PCC6803 (1001330), MalE of *E. coli* (126695) and MsmE of *S. mutans* (547925) is shown.


1 MKIKTGARILALSALTTMMFSA---SALAKIEEGKLVI-W MalE
1 MKWYKKIGLLGIVGLTSVLLAACNKSNGSQSKDDKVTIEY MsmE
1 MKFFKITTLI----ISLIVLTSCQ-GPGVNGDEDRKQVTI Slr0529
1 LKFKPGEDSR----FNWASLEEFKKGHDLKGQT-----LTI AglE
37 INGDKGYNG-LAEVGKKFEKDTG-IKVTVEHPDKLEEKFP MalE
41 FNQKKEMDATLKKIIKDFERENPKIHVKMTSVPDAGTVLK MsmE
36 LGVMIGEQQEKIEQALAPFTEATGIEVVYEGVDTFATTLP Slr0529
33 FGPWRGEDEALFKSVYAYFVEATGVELKYSSSENYEQQIV AglE
75 QVAATGDGPDII--FWAHDRFGGYAQSGLLA----EITPD MalE
81 TRVLSGDVPDVINIYPQNMDFQEWSKAGYFY----NMT-G MsmE
76 IRVDSGRAPDL-AMFPQPGLMADFAREGKLVPLGEILTPE Slr0529
73 IDTQAGSPPDV-AILPQPGLIADLAAKGLLTPLGDE-TKQ AglE
109 KAFQDKLYPFTWDAVR-YNGK-----LIAYPIAVEALSLI MalE
116 KAYLNHLKNHYANEYK-VNQK-----VYSVPLTANVSGIY MsmE
115 EMTEAY--DQAWLDLAAVDG-----TVYGVWYRASVKSLV Slr0529
111 WLLDNYAAGQSWVDLSTYNGKDGTSALYAFPYKIDVKSLV AglE
143 -YNK----DLLPNPPKTWEEIPALDKELKAKGKSALMFNL MalE
150 -YNKTKFKELGLKVPETWDEFVKLVEEIKAKKETPFALAG MsmE
148 WFNPQEFAANGYEVPGTWEEMMALSQRLIDKGKTPWCLGI Slr0529
151 WYVPENFEDAGYEVPKTMEELKALTEKIAEDGEKPWCIGL AglE
178 QEP----------YFTWPLIAADGGYA-----FKYENGKYD MalE
189 TEG----WTLNGYHQLSLISVTGSANAANKYLRFSQPN-S MsmE
188 ESGNATGWVGTDWVEDIMLRTA----SPATYDQWVAHDIP Slr0529
191 GSGGATGWPATDWVEDLMLRTQ----PAETYDKWVKNEIP AglE
204 IKDVGVDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAF MalE
224 IKTSDKILKEDMVRLNLLADDGNQQKNWKGASYNDALVAF MsmE
224 FNDRRVENA-----LDIFGEITQNEKMIYGGKVGALSTPF Slr0529
227 FTDAAVTGA-----LEEFGWFARNDAFVDGGAAAVASTDF AglE
244 NKGETAMTINGPWAWSNIDTS-KVNY----------GVT MalE
264 ANEKALMTPNGSWPCQLLNNK-IPNF----------EIG MsmE
259 GDSILGLFTDPPHCYLHRQGNFIAAFLPAD-VDDDQVDIF Slr0529
262 RDSPKGLFSSPPKCYLHHQASFIPSFFPEGKVVGEDADFF AglE
272 VLPTFKGQPSKPFVGV----LSAGINAASPNKELAKEFLE MalE
292 TFA-FPGKKTGNGITVGAGDLALSISAKTKHLKEAEKFV- MsmE
298 PLPPIEE--EYGLPILVAGDIFA-MFNDTPEARQLMAYLA Slr0529
302 YMPPYESKKELGNPVLGAGTLAM-ITKDTPAARAFIEFLK AglE
308 NYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATM MalE
330 KYMTTARAMQKYY-DVDGSPVAVKGVREDKNSPLQPLTKL MsmE
335 SSRPHEVAATLGAYISPHKNIDLNLYPDRLTRKQAEILNK Slr0529
341 TPIAHEVWMAQTSFLTPYKSVNVDVYGNPPLKKQGEILLN AglE
348 ENAQK---GEIMPNIPQMSAFW------YAVRTAVINAAS MalE
369 AFTHK---------HYVWLGQHWNSEDDLFTATTNYLMTKN MsmE
375 AEVIRFDASDMMPGAVGTGTFWSG--------MVDYIGGAD Slr0529
381 ATTFRFDGSDLMPGKIGAGAFWTG--------MVDFVGGKS AglE
379 GRQTVD--EALKDAQTRITK MalE
401 AKGLADGLNAFFNPMKADVD MsmE
408 GTQVLNTIE---RSWP---R Slr0529
414 SADVAAGVQ---KAWDSIK AglE

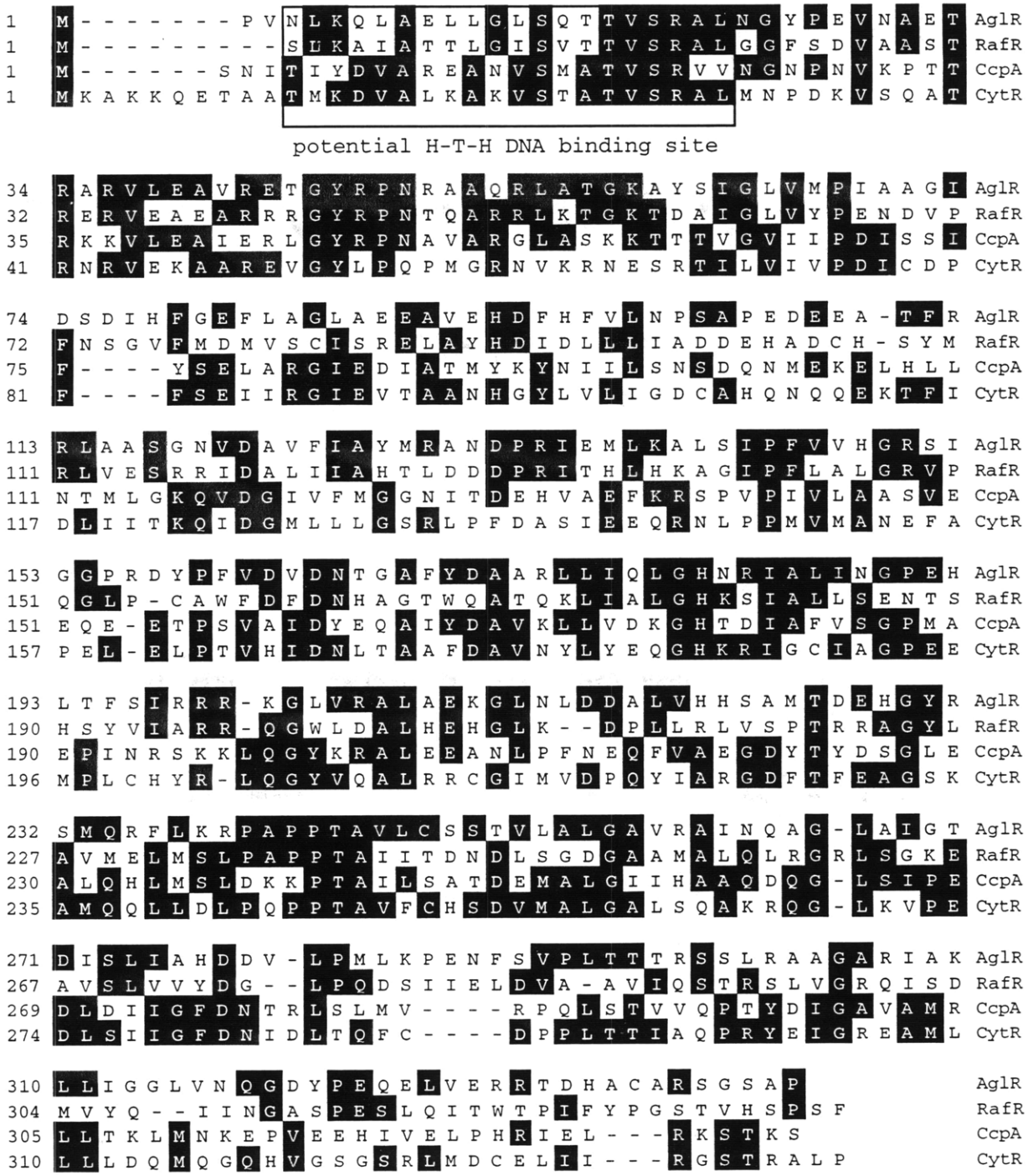


Figure 4-7

Alignment of the deduced polypeptide AglR of *R. meliloti* with proteins known to be involved in regulation of sugar catabolism genes, including RafR of *E. coli* (131825), CcpA of *B. subtilis* (1168844) and CytR of *E. coli* (118189).

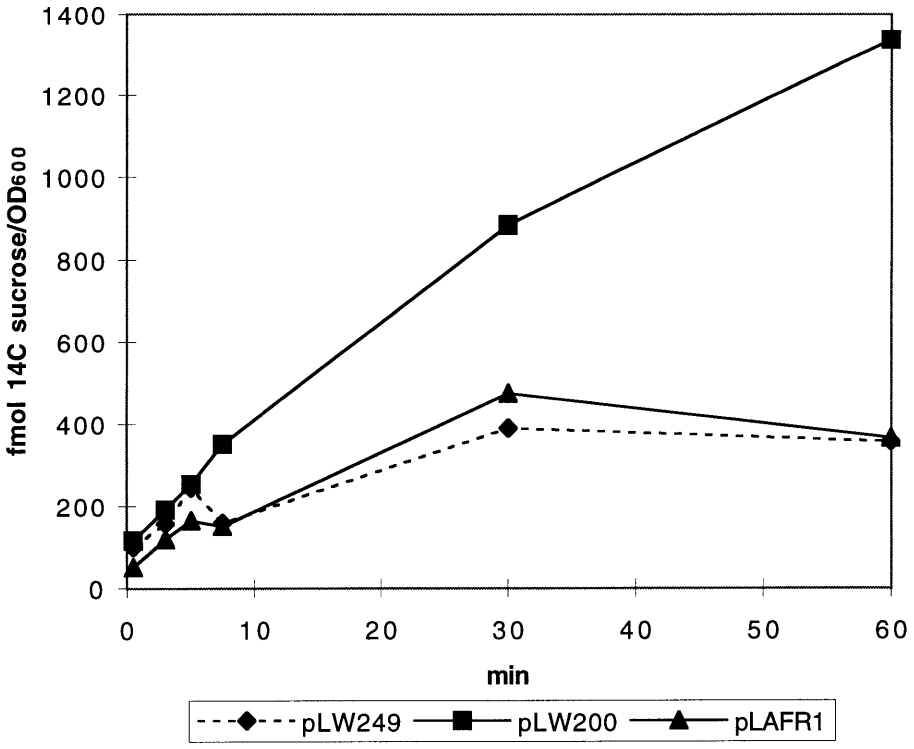


Figure 4-8.
Uptake of [U-^{14}]sucrose by *Alcaligenes eutrophus* H16 harboring pLAFR1 (▲), pLW200 (■) or pLW249 (◆).

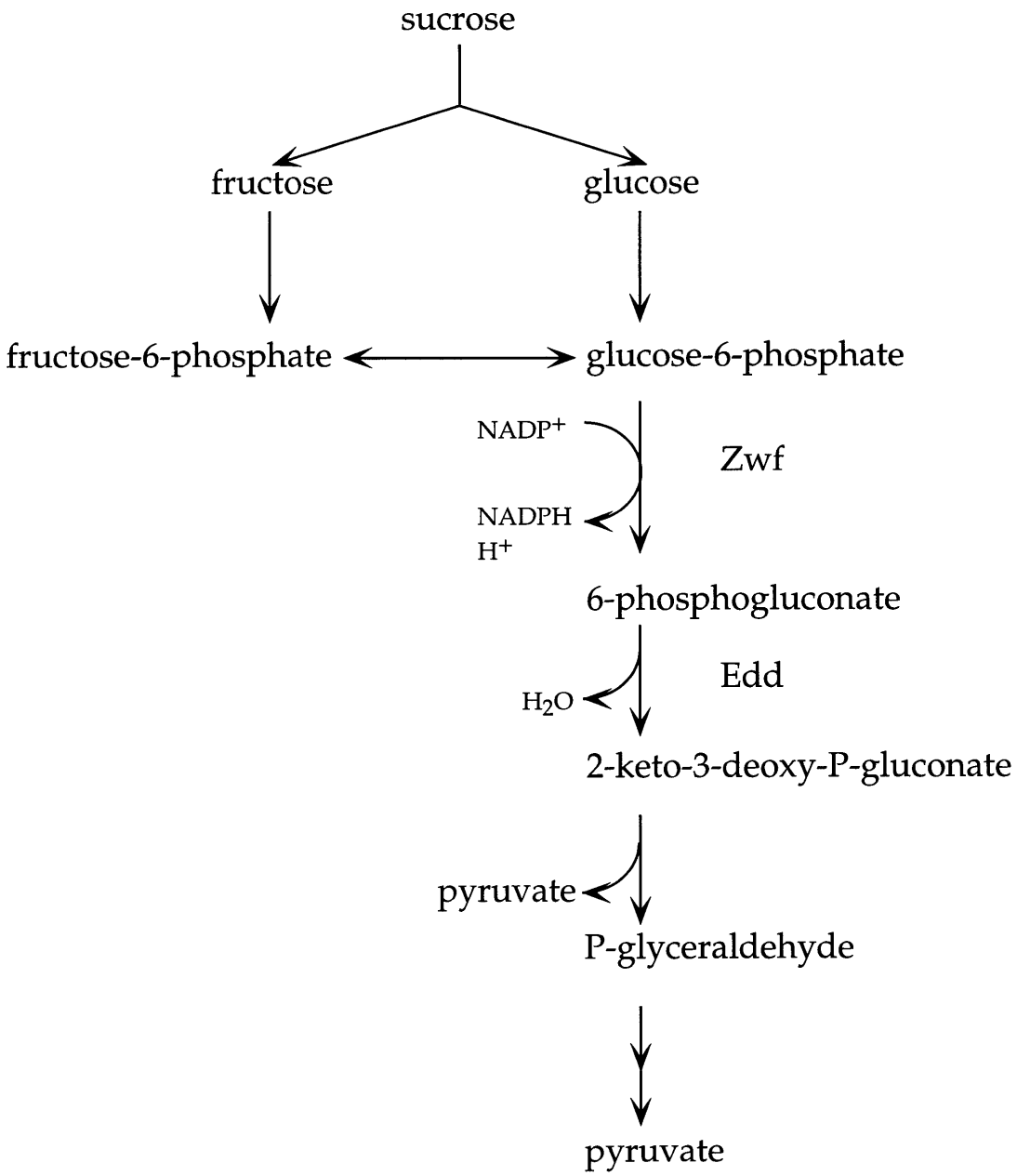
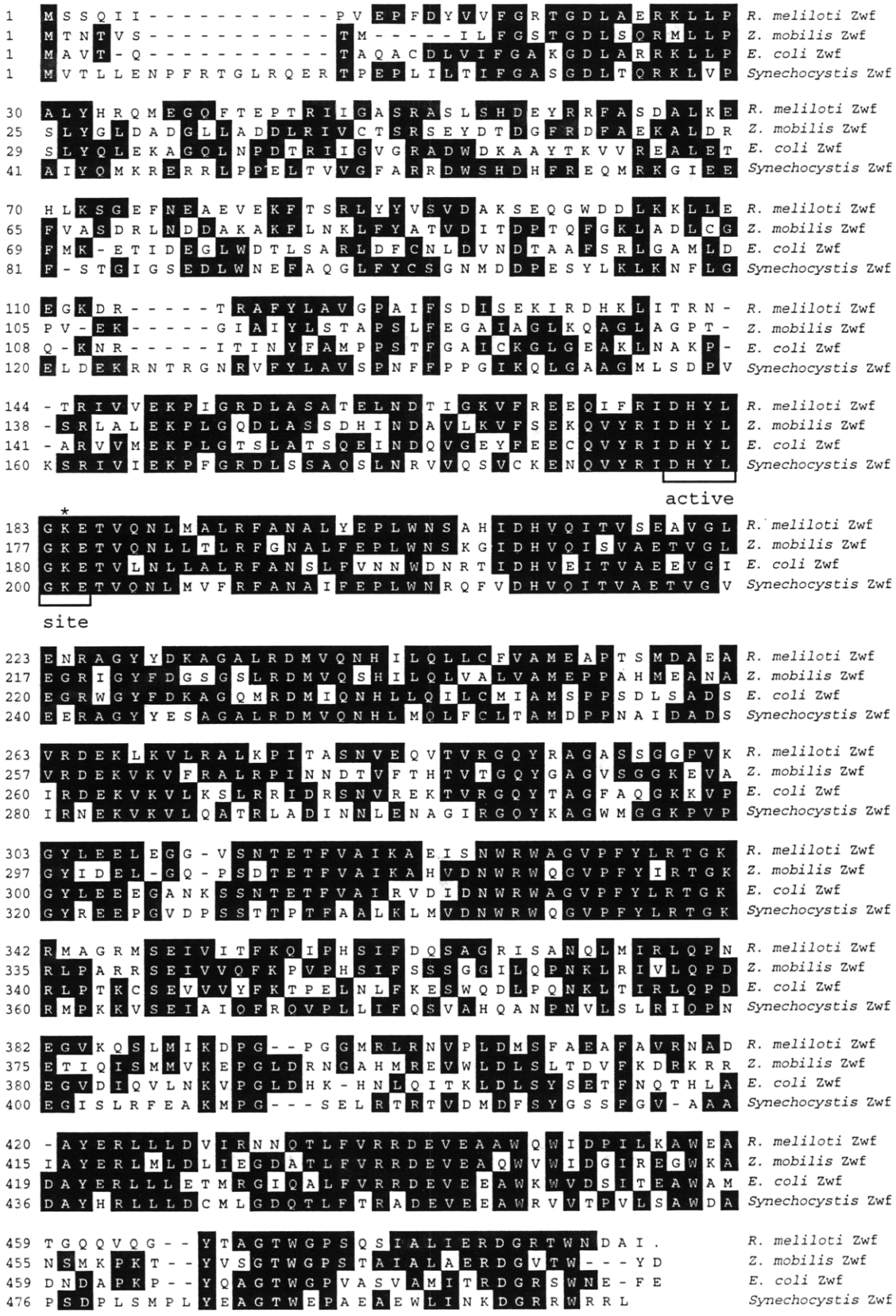


Figure 4-9.
The Entner-Doudoroff metabolic pathway for glucose utilization, showing the reactions carried out by Zwf and Edd.

Figure 4-10.
Alignment of the deduced Zwf peptides from *R. meliloti*, *Z. mobilis* (M60615), *E. coli* (2506448) and *Synechocystis* sp. strain PCC6803 (2494656). The alignment was performed using the Megalign program [Lasergene] with the Clustal algorithm. Identical residues are shaded in black. Gaps have been introduced to optimize the alignment. The PROSITE pattern conserved in glucose-6-phosphate dehydrogenase active sites is boxed. The presumed active site lysine is indicated with an asterisk. Accession numbers are indicated in parentheses.


1 MSSQII----------PVEPFDYVVFGRTGDLAERKLLP R. meliloti Zwf
1 MTNTVS----------TM------ILFGSTGDLSQRMLLP Z. mobilis Zwf
1 MAVT-Q----------TAQACDLVIFGAKGDLARRKLLP E. coli Zwf
1 MVTLLENPFRTGLRQERTPEPLILTIFGASGDLTQRKLVP Synechocystis Zwf
30 ALYHRQMEGQFTEPTRIIGASRASLSHDEYRRFASDALKE R. meliloti Zwf
25 SLYGLDADGLLADDLRIVCTSRSEYDTDGFRDFAEKALDR Z. mobilis Zwf
29 SLYQLEKAGQLNPDTRIIGVGRADWDKAAYTKVVREALET E. coli Zwf
41 AIYQMKRERRLPPELTVVGFARRDWSHDHFREQMRKGIEE Synechocystis Zwf
70 HLKSGEFNEAEVEKFTSRLYYVSVDAKSEQGWDDLKKLLE R. meliloti Zwf
65 FVASDRLNDDAKAKFLNKLFYATVDITDPTQFGKLADLCG Z. mobilis Zwf
69 FMK-ETIDEGLWDTLSARLDFCNLDVNDTAAFSRLGAMLD E. coli Zwf
81 F-STGIGSEDLWNEFAQGLFYCSGNMDDPESYLKLKNFLG Synechocystis Zwf
110 EGKDR-----TRAFYLAVGPAIFSDISEKIRDHKLITRN- R. meliloti Zwf
105 PV-EK-----GIAIYLSTAPSLFEGAIAGLKQAGLAGPT- Z. mobilis Zwf
108 Q-KNR-----ITINYFAMPPSTFGAICKGLGEAKLNAKP- E. coli Zwf
120 ELDEKRNTRGNRVFYLAVSPNFFPPGIKQLGAAGMLSDPV Synechocystis Zwf
144 -TRIVVEKPIGRDLASATELNDTIGKVFREEQIFRIDHYL R. meliloti Zwf
138 -SRLALEKPLGQDLASSDHINDAVLKVFSEKQVYRIDHYL Z. mobilis Zwf
141 -ARVVMEKPLGTSLATSQEINDQVGEYFEECQVYRIDHYL E. coli Zwf
160 KSRIVIEKPFGRDLSSAQSLNRVVQSVCKENQVYRIDHYL Synechocystis Zwf
active
*
183 GKETVQNLMALRFANALYEPLWNSAHIDHVQITVSEAVGL R. meliloti Zwf
177 GKETVQNLLTLRFGNALFEPLWNSKGIDHVQISVAETVGL Z. mobilis Zwf
180 GKETVLNLLALRFANSLFVNNWDNRTIDHVEITVAEEVGI E. coli Zwf
200 GKETVQNLMVFRFANAIFEPLWNRQFVDHVQITVAETVGV Synechocystis Zwf
site
223 ENRAGYYDKAGALRDMVQNHILQLLCFVAMEAPTSMDAEA R. meliloti Zwf
217 EGRIGYFDGSGSLRDMVQSHILQLVALVAMEPPAHMEANA Z. mobilis Zwf
220 EGRWGYFDKAGQMRDMIQNHLLQILCMIAMSPPSDLSADS E. coli Zwf
240 EERAGYYESAGALRDMVQNHLMQLFCLTAMDPPNAIDADS Synechocystis Zwf
263 VRDEKLKVLRALKPITASNVEQVTVRGQYRAGASSGGPVK R. meliloti Zwf
257 VRDEKVKVFRALRPINNDTVFTHTVTGQYGAGVSGGKEVA Z. mobilis Zwf
260 IRDEKVKVLKSLRRIDRSNVREKTVRGQYTAGFAQGKKVP E. coli Zwf
280 IRNEKVKVLQATRLADINNLENAGIRGQYKAGWMGGKPVP Synechocystis Zwf
303 GYLEELEGG-VSNTETFVAIKAEISNWRWAGVPFYLRTGK R. meliloti Zwf
297 GYIDEL-GQ-PSDTETFVAIKAHVDNWRWQGVPFYIRTGK Z. mobilis Zwf
300 GYLEEEGANKSSNTETFVAIRVDIDNWRWAGVPFYLRTGK E. coli Zwf
320 GYREEPGVDPSSTTPTFAALKLMVDNWRWQGVPFYLRTGK Synechocystis Zwf
342 RMAGRMSEIVITFKQIPHSIFDQSAGRISANQLMIRLQPN R. meliloti Zwf
335 RLPARRSEIVVQFKPVPHSIFSSSGGILQPNKLRIVLQPD Z. mobilis Zwf
340 RLPTKCSEVVVYFKTPELNLFKESWQDLPQNKLTIRLQPD E. coli Zwf
360 RMPKKVSEIAIQFRQVPLLIFQSVAHQANPNVLSLRIQPN Synechocystis Zwf
382 EGVKQSLMIKDPG--PGGMRLRNVPLDMSFAEAFAVRNAD R. meliloti Zwf
375 ETIQISMMVKEPGLDRNGAHMREVWLDLSLTDVFKDRKRR Z. mobilis Zwf
380 EGVDIQVLNKVPGLDHK-HNLQITKLDLSYSETFNQTHLA E. coli Zwf
400 EGISLRFEAKMPG---SELRTRTVDMDFSYGSSFGV-AAA Synechocystis Zwf
420 -AYERLLLDVIRNNQTLFVRRDEVEAAWQWIDPILKAWEA R. meliloti Zwf
415 IAYERLMLDLIEGDATLFVRRDEVEAQWVWIDGIREGWKA Z. mobilis Zwf
419 DAYERLLLETMRGIQALFVRRDEVEEAWKWVDSITEAWAM E. coli Zwf
436 DAYHRLLLDCMLGDQTLFTRADEVEEAWRVVTPVLSAWDA Synechocystis Zwf
459 TGQQVQG--YTAGTWGPSQSIALIERDGRTWNDAI. R. meliloti Zwf
455 NSMKPKT--YVSGTWGPSTAIALAERDGVTW---YD Z. mobilis Zwf
459 DNDAPKP--YQAGTWGPVASVAMITRDGRSWNE-FE E. coli Zwf
476 PSDPLSMPLYEAGTWEPAEAEWLINKDGRRWRRL Synechocystis Zwf

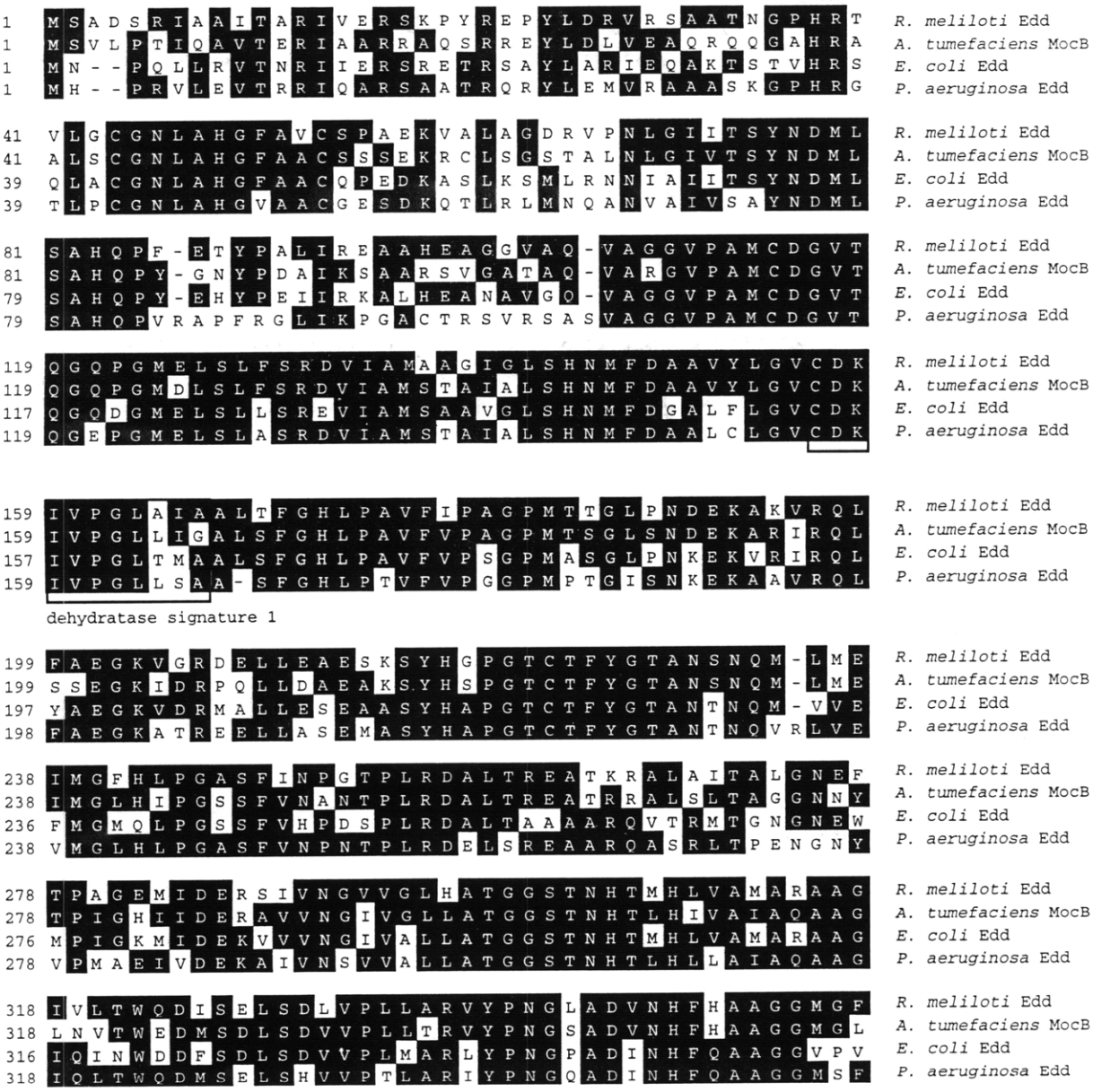
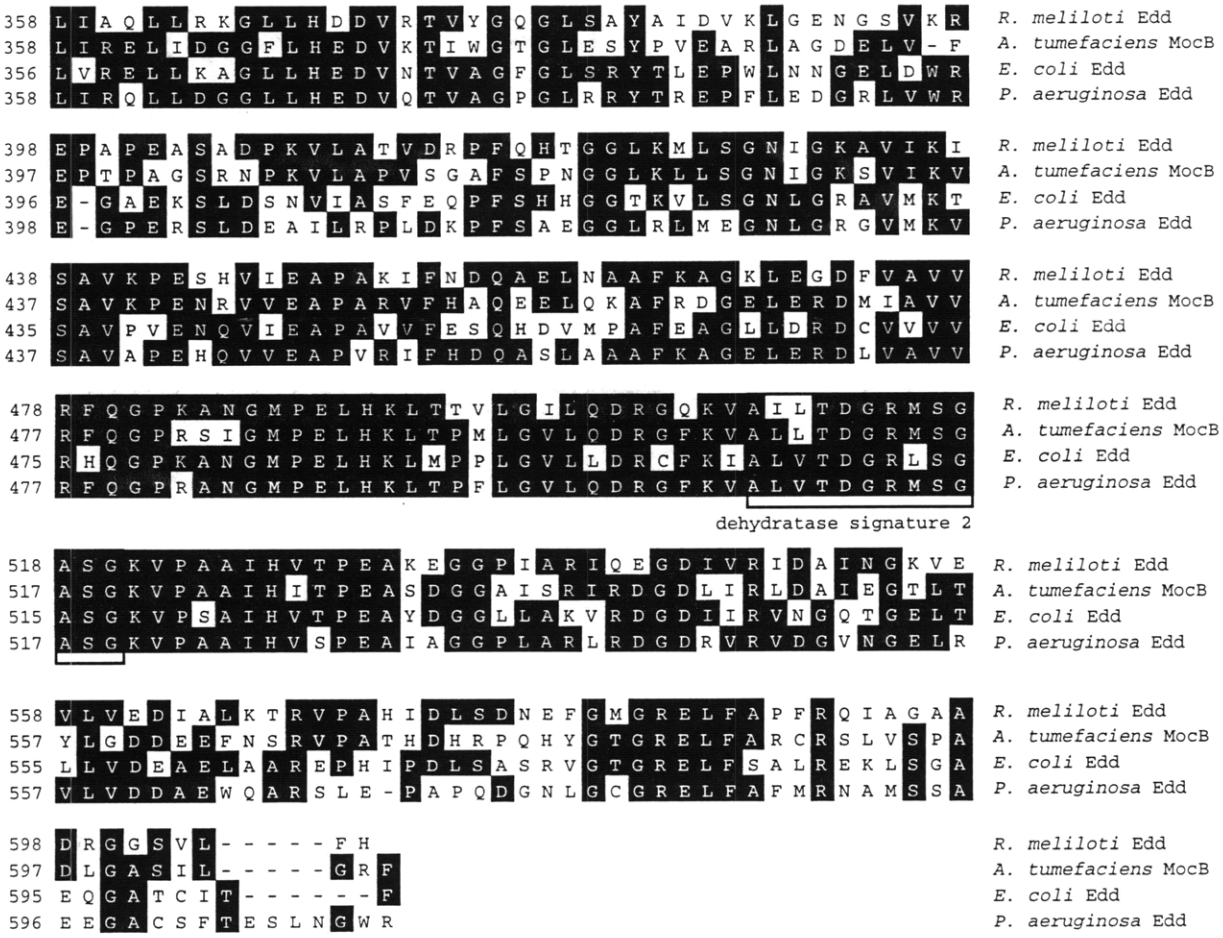


Figure 4-11.

Alignment of the deduced dehydratase proteins Edd of *R. meliloti*, MocB of *A. tumefaciens* (U19620), Edd of *E. coli* (461986) and Edd of *P. aeruiginosa* (1169470). The alignment was performed using the Clustal algorithm of Megalign [Lasergene]. The PROSITE (Bairoch, 1992) patterns dehydratase signature 1 and dehydratase signature 2 are boxed. Identical residues are shaded in black. Gaps have been introduced to optimize the alignment. Accession numbers are indicated in parentheses.


358 LIAQLLRKGLLHDDVRTVYGQGLSAYAIDVKLGENGSVKR R. meliloti Edd
358 LIRELIDGGFLHEDVKTIWGTGLESYPVEARLAGDELV-F A. tumefaciens MocB
356 LVRELLKAGLLHEDVNTVAGFGLSRYTLEPWLNNGELDWR E. coli Edd
358 LIRQLLDGGLLHEDVQTVAGPGLRRYTREPFLEDGRLVWR P. aeruginosa Edd
398 EPAPEASADPKVLATVDRPFQHTGGLKMLSGNIGKAVIKI R. meliloti Edd
397 EPTPAGSRNPKVLAPVSGAFSPNGGLKLLSGNIGKSVIKV A. tumefaciens MocB
396 E-GAEKSLDSNVIASFEQPFSHHGGTKVLSGNLGRAVMKT E. coli Edd
398 E-GPERSLDEAILRPLDKPFSAEGGLRLMEGNLGRGVMKV P. aeruginosa Edd
438 SAVKPESHVIEAPAKIFNDQAELNAAFKAGKLEGDFVAVV R. meliloti Edd
437 SAVKPENRVVEAPARVFHAQEELQKAFRDGELERDMIAVV A. tumefaciens MocB
435 SAVPVENQVIEAPAVVFESQHDVMPAFEAGLLDRDCVVVV E. coli Edd
437 SAVAPEHQVVEAPVRIFHDQASLAAAFKAGELERDLVAVV P. aeruginosa Edd
478 RFQGPKANGMPELHKLTTVLGILQDRGQKVAILTDGRMSG R. meliloti Edd
477 RFQGPRSIGMPELHKLTPMLGVLQDRGFKVALLTDGRMSG A. tumefaciens MocB
475 RHQGPKANGMPELHKLMPPLGVLLDRCFKIALVTDGRLSG E. coli Edd
477 RFQGPRANGMPELHKLTPFLGVLQDRGFKVALVTDGRMSG P. aeruginosa Edd
dehydratase signature 2
518 ASGKVPAAIHVTPEAKEGGPIARIQEGDIVRIDAINGKVE R. meliloti Edd
517 ASGKVPAAIHITPEASDGGAISRIRDGDLIRLDAIEGTLT A. tumefaciens MocB
515 ASGKVPSAIHVTPEAYDGGLLAKVRDGDIIRVNGQTGELT E. coli Edd
517 ASGKVPAAIHVSPEAIAGGPLARLRDGDRVRVDGVNGELR P. aeruginosa Edd
558 VLVEDIALKTRVPAHIDLSDNEFGMGRELFAPFRQIAGAA R. meliloti Edd
557 YLGDDEEFNSRVPATHDHRPQHYGTGRELFARCRSLVSPA A. tumefaciens MocB
555 LLVDEAELAARVPHIPDLSASRVGTGRELFSALREKLSGA E. coli Edd
557 VLVDDAEWQARSLE-PAPQDGNLGCGRELFAFMRNAMSSA P. aeruginosa Edd
598 DRGGSVL-----FH R. meliloti Edd
597 DLGASIL-----GRF A. tumefaciens MocB
595 EQGATCIT------F E. coli Edd
596 EEGACSFTESLNGWR P. aeruginosa Edd

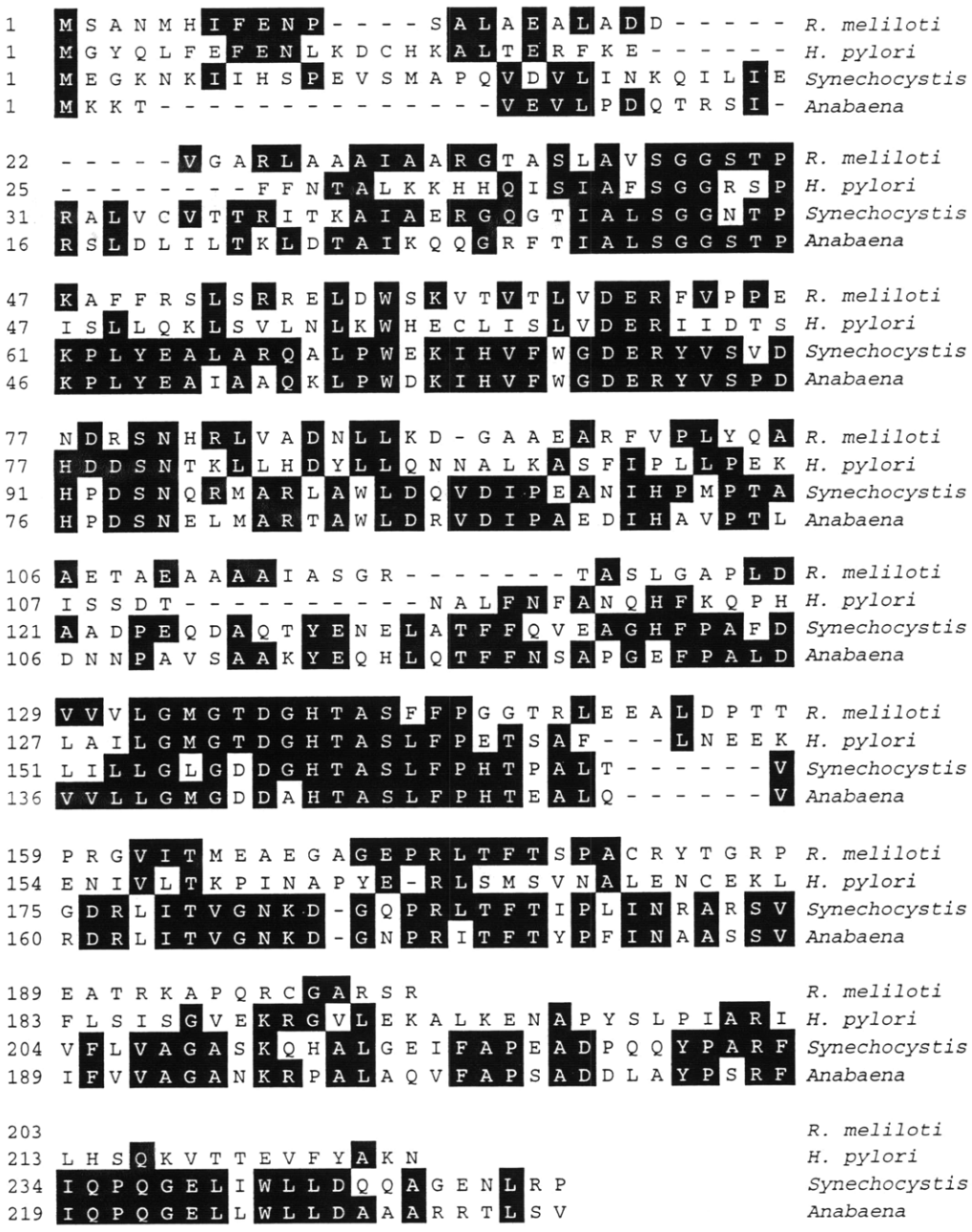


Figure 4-12.

Alignment of the deduced DevB peptides from *R. meliloti*, *H. pylori* (AE000616), *Synechocystis* sp. strain PCC6803 (D90916) and *Anabaena* sp. strain PCC7120 (U14553). The sequences were aligned using the Clustal algorithm of Megalign on a Power Macintosh computer. Identical residues are shaded with black. Gaps have been introduced to optimize the alignment. Genbank accession numbers are included in parentheses.

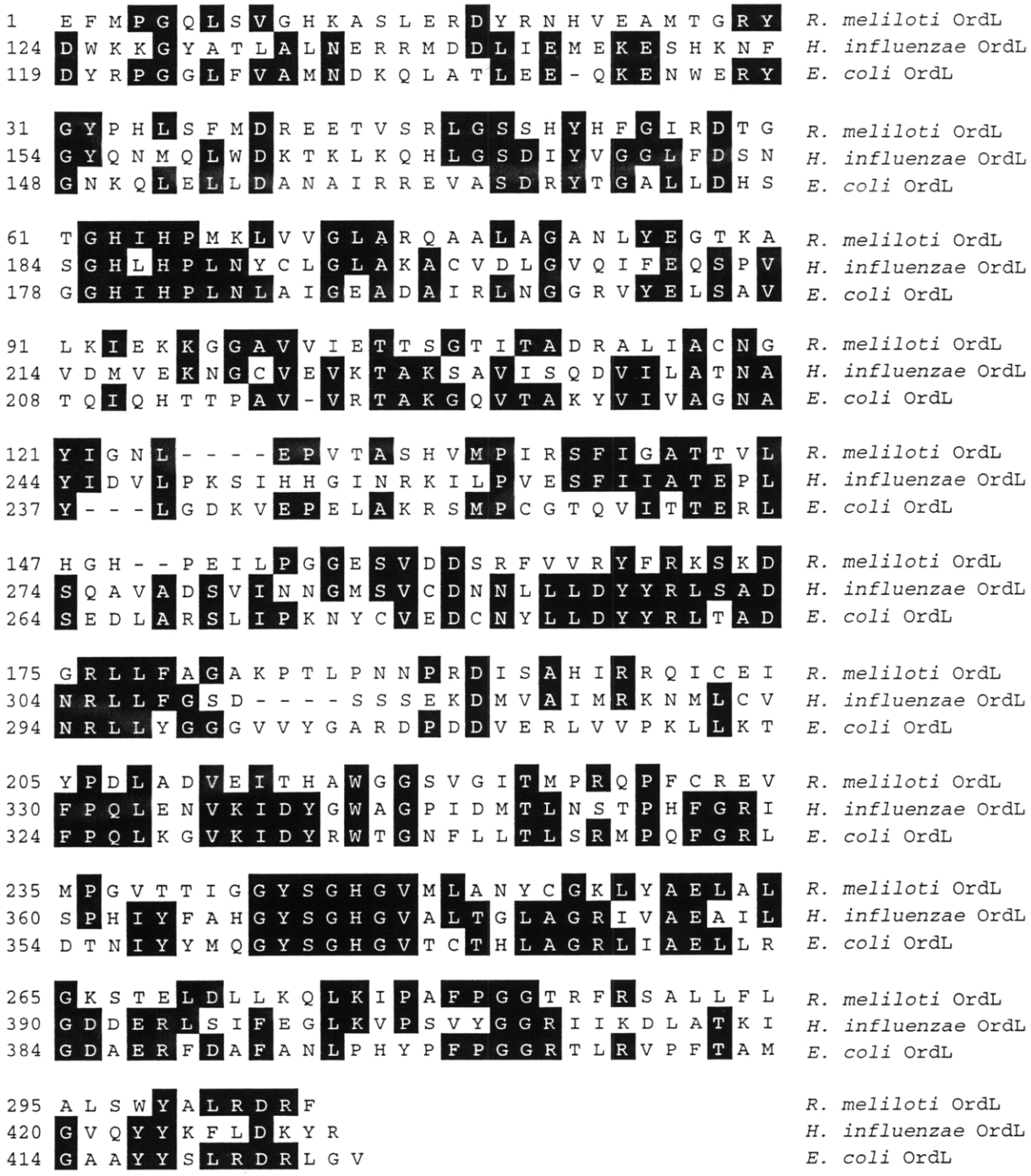


Figure 4-13.

Alignment of the deduced peptide from *R. meliloti* partial open reading frame *ordL* and the C-terminal portions of *H. influenzae* OrdL (P44732) and *E. coli* OrdL (P37906). The alignment was performed using the Clustal algorithm of Megalign [Lasergene]. Identical residues are shaded in black. Gaps have been introduced to optimize the alignment. Accession numbers are indicated in parentheses.

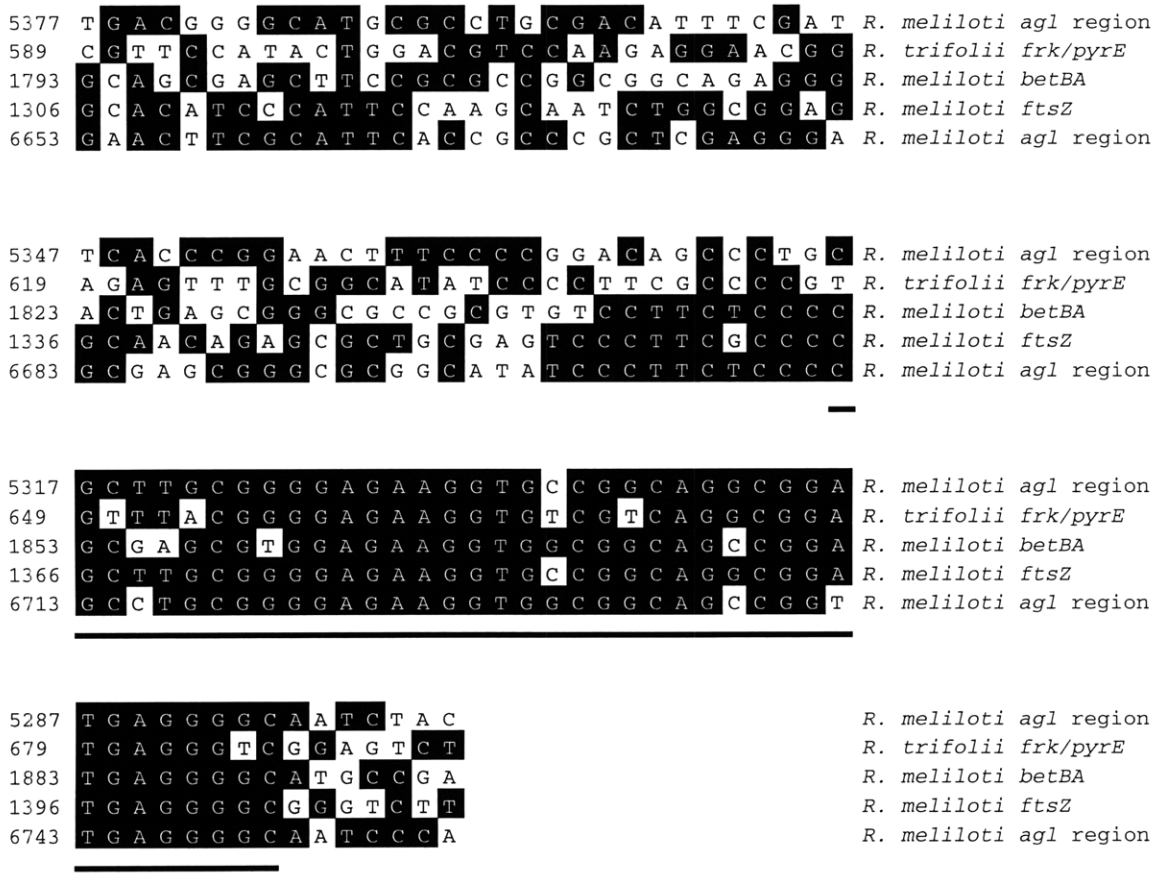


Figure 4-14.

Alignment of the intergenic regions of *R. meliloti agl* region, *R. meliloti ftsZ* (L25440), *R. trifolii frk/pyrE* (U08434) and *R. meliloti betBA* (U39940). The strongly conserved motif which flanks *tnp* in the *R. meliloti agl* region is underlined.

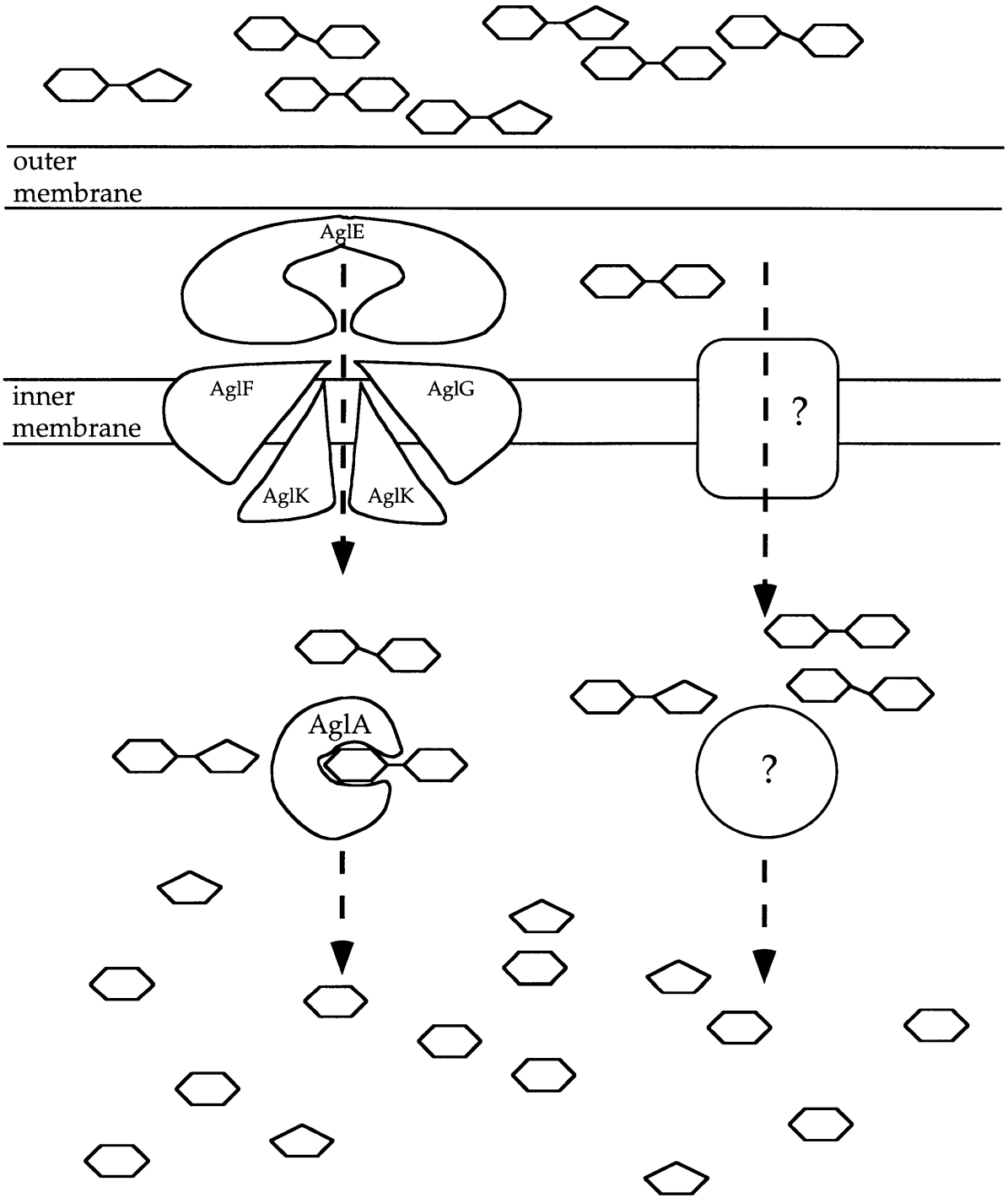


Figure 4-15.
Model for *R. meliloti* alpha glucoside utilization.

Chapter 5

Concluding Remarks

In this thesis we have explored issues of carbon metabolism by *Rhizobium meliloti* and identified several genes involved in carbon metabolism. The observation that PHB granules, which are present in *R. meliloti* in the infection thread, disappear soon after release into the symbiosome membrane prompted the experiment to determine whether disruption of the gene encoding PHB synthase would have an effect on symbiosis. We cloned the *phbC* gene and isolated mutants that do not produce PHB but found that they are not compromised in their effectiveness in the symbiosis. This result is consistent with the correlation between low levels of PHB and high activity of nitrogenase (Romanov et al., 1974; Tombolini and Nuti, 1989). Perhaps *R. meliloti* has evolved the ability to turn off PHB biosynthesis *in planta* in order to maximize its nitrogenase activity.

We also explored sucrose metabolism by *R. meliloti* and found that conditions used to screen for genes which allow a heterologous host to utilize sucrose can greatly affect the outcome of the experiment. Indeed, the alcohol dehydrogenase gene that was identified in a screen for *R. meliloti* genomic clones that improve the growth of an *A. eutrophus phbC* mutant on sucrose was not isolated when the screen was repeated in a wild type *A. eutrophus* background. We showed that *R. meliloti adhA* is expressed in *A. eutrophus* and offered a model for why *R. meliloti adhA* may benefit a *A. eutrophus phbC* strain.

The most significant findings from this work are the characterization of the *agl* genes for α-glucoside utilization in *R. meliloti*. We found that, unlike most bacteria, *R. meliloti* does not use the phosphotransferase system for the utilization of sucrose but instead appears to employ a periplasmic binding protein dependent transport system. AglE and the *Synechocystis* homologue Slr0529 may represent a new class of periplasmic binding proteins with substrate specificity for α-glucosides.

Biochemical characterization of Agl proteins

Although we have shown conclusively that the *agl* genes are involved in α-glucoside utilization and demonstrated that the *agl* genes promote uptake of ^{14}C-sucrose, much of our model is based solely on the homologies of deduced polypeptides. Further biochemical characterization of the Agl proteins will be required to prove or refute the functions we have assigned to them.

We have suggested that *aglE* may encode a periplasmic substrate binding protein, based on the limited homology with *E. coli* MalE and *S. mutans* MsmE (see Figure 4-6), and on the conserved gene order in operons encoding binding-protein dependent transport systems (Boos and Lucht, 1996). To vigorously test that hypothesis, it would be necessary to show both that AglE is found in the periplasm and that it binds α-glucosides. As suggested in Chapter 4, one straightforward experiment involves isolating periplasmic proteins using an osmotic shock protocol and then isolating the proteins that bind to an amylose-agarose resin. If one of the two prominent periplasmic maltose-binding proteins is absent in protein preparations from an *aglE* mutant, it would argue strongly that AglE is found in the periplasm. The identification of a particular protein that binds to amylose-agarose resin as AglE could be confirmed by subjecting the protein to N-terminal protein sequencing and comparing the data with the predicted amino acid sequence of AglE. A more direct method to show maltose binding could involve overexpressing *aglE* in *E. coli* and demonstrating that the overproduced protein binds to amylose resin.

To test the hypotheses that AglE is a periplasmic protein and that AglF and AglG are inner membrane proteins, Tn*phoA* fusions could be constructed to the *aglE*, *aglF* and *aglG* genes. This derivative of Tn*5* makes fusions to the reporter alkaline phosphatase (Manoil and Beckwith, 1985), which is active in the periplasm but not

in the cytoplasm. If strains carrying Tn*phoA* insertions in these genes produce blue colonies on medium containing the chromogenic substrate X-P, it would indicate that the fusion protein is in the periplasm.

Our observations that *R. meliloti aglA* mutants are not as severely affected for growth on α-glucosides as *aglE* or *aglF* mutants suggests both that these strains still have sucrose hydrolysis activity and that α-glucosides are still transported in these strains. ^{14}C-sucrose uptake assays can be carried out in an *A. eutrophus* strain carrying a derivative of pLW200 that has a Tn*5* insertion in the *aglA* gene to test the latter assumption. To test the former, protein preparations from wild type and *aglA* strains of *R. meliloti*, or *A. eutrophus* harboring pLW200 or its mutagenized derivatives, could be tested *in vitro* for the ability to produce reducing ends from the substrate sucrose. Such studies could take into account the observation by Hoelzle and Streeter that α-glucosidase activity in *R. meliloti* is stimulated by potassium, rubidium and ammonium ions (Hoelzle and Streeter, 1990). Once conditions are established which are optimized for sucrose hydrolysis activity, it would be practical to examine α-glucosidase activity in non-denaturing polyacrylamide gels (Manchenko, 1994). The advantage of this method is that it makes it possible to determine the substrate specificity of glycosyl hydrolase activities by examining, for example, whether proteins which cleave sucrose and maltose are also active on isomaltose.

Our model predicts that AglR is a transcriptional regulator and that it may have a helix-turn-helix DNA binding motif. This hypothesis could be tested by performing gel shift assays using purified AglR protein and DNA from the presumptive promoter region between *aglE* and *aglR*. There exists a direct repeat just upstream of the postulated start site of *aglE* which could act as a binding site for a regulatory protein.

A possible role for *zwf* in synthesis of succinoglycan

The acidic exopolysaccharide succinoglycan fluoresces under UV light after binding the laundry whitener Calcofluor. This property has been exploited to screen for *R. meliloti* mutants which fail to produce succinoglycan and are therefore dark on Calcofluor plates (Leigh et al., 1985; Long et al., 1988). Several additional *R. meliloti* mutants have been identified that form brighter than wild type colonies on LB medium containing Calcofluor (Leigh and Walker, Unpublished results). These mutants have not been extensively characterized and are distinct from the strains which overproduce succinoglycan (Doherty et al., 1988). It seems possible that a *zwf* mutant might accumulate glucose-6-phosphate which could be converted into UDP-glucose, a precursor for succinoglycan biosynthesis. If so, some mutants which produce bright colonies on medium containing Calcofluor may carry disruptions in *zwf*. This hypothesis could be tested by determining whether the Calcofluor-bright mutants fail to grow on sucrose, as would be predicted for a glucose-6-phosphate dehydrogenase mutant (Cerveñansky and Arias, 1984) and whether diminution of the bright phenotype is seen when pLW200 is introduced into the Calcofluor-bright mutants. Although this model is purely speculative and the cell tends to tightly regulate levels of UDP-glucose, a *zwf* mutant of *Salmonella typhimurium* was found to accumulate TDP-rhamnose, which was incorporated into a polymer produced by that strain (Wright, Personal communication).

Mutagenesis of *agl* mutant strains

Because our model predicts that *R. meliloti* contains additional genes involved in sucrose uptake or hydrolysis, we mutagenized *aglA112*::Tn*5-233* (Rm9631) and *aglE49*::Tn*5-233* (Rm9632) with Tn*5* and assessed the ability of the double mutants to

utilize carbon sources by replica plating onto minimal plates containing sucrose, maltose, trehalose or succinate. After screening approximately 7600 double mutants, we identified four *aglA112*::Tn*5-233* ΩTn*5* strains and one *aglE49*::Tn*5-233* ΩTn*5* strain which are unable to grow on sucrose, maltose or trehalose but still able to grow on succinate. Each of the Tn*5* insertions from the double mutants was transduced into Rm1021 and Rm9631 or Rm9632 to ensure a clean genetic background. The Tn*5* insertions in these strains may define new loci involved in sucrose uptake or hydrolysis. Further characterization of these strains should involve the isolation of *R. meliloti* genomic clones which complement the mutation(s).

During the mutagenesis of *aglE*::Tn*5-233* strain Rm9632, which was accomplished by transducing the strain with a bacteriophage ϕM12 lysate that had been grown on a pool of Rm1021 derivatives containing random Tn*5* insertions, approximately 1% of the transductants were observed to grow more strongly on sucrose plates than the parent *aglE* strain. These strains proved to have transduced in a wild type copy of *aglE*, because they are Nm^R but have lost the Gm^R/Sp^R marker from Tn*5-233*. Since the mutations in these strains are transductionally linked to *aglE*, their Tn*5* insertions are predicted to map within approximately 5 kb of *aglE*. Some of the Tn*5* insertions in these strains may lie within loci such as *devB* for which we currently have only sequence homology information. It would be straightforward to determine which of these Tn*5* insertions lie within the region that has been sequenced by performing a Southern blot on *Eco*RI-digested genomic DNA from these mutants, using labeled pLW200 as a probe. Because Tn*5* does not contain any *Eco*RI sites, Tn*5* insertions which lie within one of the 10 *Eco*RI fragments present in pLW200 will be easily detected by the decreased mobility of the hybridizing band.

If genes encoding the proposed additional glycanase(s) are not found in a screen for derivatives of *aglA* which cannot utilize α-glucosides, it may be possible to identify them by hybridization. The active site residues of α-glucosidases are very conserved (Svensson, 1994; Janecek et al., 1997), and an alignment of deduced protein sequences could be used to design degenerate oligonucleotides for use as a probe or as PCR primers. Preliminary evidence suggests that such an approach would be fruitful. A weakly hybridizing band of approximately 10 kb is seen when *R. meliloti* genomic DNA is probed with the insert from pLW201, which contains the C-terminal half of *aglG* and the majority of *aglA* (data not shown). This band may represent a gene encoding a glycanase or permease.

REFERENCES

Alfano, J. R., and Kahn, M. L. (1993). Isolation and characterization of a gene coding for a novel aspartate aminotransferase from *Rhizobium meliloti*. *J. Bacteriol.* **175** : 4186-4196.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215** : 403-410.

Anderson, A. J., and Dawes, E. A. (1990). Occurrence, metabolism, metabolic role, and industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* **54 (4):** 450-472.

Antoniw, L. D., and Sprent, J. I. (1978). Primary metabolites of *Phaseolus vulgaris* nodules. *Phytochemistry* **17** : 675.

Arias, A., and Cerveñansky, C. (1986). Galactose metabolism in *Rhizobium meliloti* L5-30. *J. Bacteriol.* **167 (3):** 1092-1094.

Arias, A., Cerveñansky, C., Gardiol, A., and Martinez-Drets, G. (1979). Phosphoglucose isomerase mutant of *Rhizobium meliloti*. *J. Bacteriol.* **137 (1):** 409-414.

Arnaud, M., Débarbouilé, M., Rapoport, G., Zagorec, M., and Saier, M. H., Jr. (1996). *In vitro* reconstitution of transcriptional antitermination by the SacT and SacY proteins of *Bacillus subtilis*. *J. Biol. Chem.* **271** : 18966-18972.

Aslanidis, C., Schmid, K., and Schmitt, R. (1989). Nucleotide sequences and operon structure of plasmid-borne genes mediating uptake and utilization of raffinose in *Escherichia coli*. *J. Bacteriol.* **171 (12):** 6753-6763.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1995). Current Protocols in Molecular Biology, K. Janssen, ed. (New York: Current Protocols).

Bairoch, A. (1992). PROSITE: A dictionary of sites and patterns in proteins. *Nucleic Acids Res.* **20** : 2013-2018.

Bakkers, J., Semino, C. E., Stroband, H., Kijne, J. W., Robbins, P. W., and Spaink, H. P. (1997). An important developmental role for oligosaccharides during early embryogenesis of cyprinid fish. *Proc. Natl. Acad. Sci. USA* **94 (15):** 7982-7986.

Bardin, S., Dan, S., Østerås, M., and Finan, T. M. (1996). A phosphate transport system is required for symbiotic nitrogen fixation by *Rhizobium meliloti*. *J. Bacteriol.* **178 (15):** 4540-4547.

Barnell, W. O., Yi, K. C., and Conway, T. (1990). Sequence and genetic organization of a *Zymomonas mobilis* gene cluster that encodes several enzymes of glucose metabolism. *J. Bacteriol.* **172 :** 7227-7240.

Bauer, C. C. (1994). Isolation and characterization of genes involved in nitrogen fixation and heterocyst differentiation in *Anabaena* sp. Strain PCC7120. Ph. D. thesis. (Chicago, Ill.: The University of Chicago).

Bavoil, P., Hofnung, M., and Nikaido, H. (1980). Identification of a cytoplasmic membrane-associated component of the maltose transport system of *Escherichia coli. J. Biol. Chem.* **255 (18):** 8366-8369.

Becker, A., Kleickmann, A., Keller, M., Arnold, W., and Pühler, A. (1993). Identification and analysis of the *Rhizobium meliloti exoAMNOP* genes involved in exopolysaccharide biosynthesis and mapping of promoters located on the *exoHKLAMONP* fragment. *Mol. Gen. Genet.* **241 :** 367-379.

Becker, A., Rüberg, S., Küster, H., Roxlau, A. A., Keller, M., Ivashina, T., Cheng, H.-P., Walker, G. C., and Pühler, A. (1997). The 32-kilobase *exp* gene cluster of *Rhizobium meliloti* directing the biosynthesis of galactoglucan: genetic organization and properties of the encoded gene products. *J. Bacteriol.* **179 (4):** 1375-1384.

Bedouelle, H., and Duplay, P. (1988). Production in *Escherichia coli* and one-step purification of bi-functional hybrid proteins which bind maltose. *Eur. J. Biochem.* **171 :** 541-549.

Bergersen, F. J., Peoples, M. B., and Turner, G. L. (1991). A role for poly-β-hydroxybutyrate in bacteroids of soybean root nodules. *Proc. R. Soc. Lond. B* **245 :** 59-64.

Beringer, J. E., Benyon, J. L., Buchanan-Wollaston, A. V., and Johnston, A. W. B. (1978). Transfer of the drug-resistance transposon Tn*5* to *Rhizobium. Nature* **276 :** 633-634.

Bockman, J., Heuel, H., and Lengeler, J. W. (1992). Characterization of a chomosomally encoded, non-PTS metabolic pathway for sucrose utilization in *Escherichia coli* EC3132. *Mol. Gen. Genet.* **235 :** 22-32.

Boivin, C., Camut, S., Malpica, C. A., Truchet, G., and Rosenberg, C. (1990). *Rhizobium meliloti* genes encoding the catabolism of trigonelline are induced under symbiotic conditions. *Plant Cell* **2 :** 1157-1170.

Bolton, E., Higgisson, B., Harrington, A., and O'Gara, F. (1986). Dicarboxylic acid transport in *Rhizobium meliloti*: Isolation of mutants and cloning of dicarboxylic acid transport genes. *Arch. Microbiol.* **144 :** 142-146.

Boos, W., and Lucht, J. M. (1996). Periplasmic binding protein-dependent ABC transporters. In *Escherichia coli and Salmonella: cellular and molecular biology.,*

F. C. Neidhardt, R. I. Curtiss, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter and H. E. Umbarger, eds. (Washington, D. C.: ASM Press), pp. 1175-1209.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72 :** 248-254.

Brewin, N. J. (1991). Development of the legume root nodule. *Ann. Rev. Cell Biol.* **7 :** 191-226.

Brueckner, R., Wagner, E., and Goetz, F. (1993). Characterization of a sucrase gene from *Staphylococcus xylosus. J. Bacteriol.* **175 (3):** 851-857.

Burdon, K. L. (1946). Fatty Material in Bacteria and Fungi Revealed by Staining Dried, Fixed Slide Preparations. *J. Bacteriol.* **52 :** 665-678.

Caetano-Anolles, G., Crist-Estes, D. K., and Bauer, W. D. (1988). Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin requires functional nodulation genes. *J. Bacteriol.* **170 :** 3164-3169.

Cannio, R., Rossi, M., and Bartolucci, S. (1994). A few amino acid substitutions are responsible for the higher thermostability of a novel NAD(+)-dependent bacillar alcohol dehydrogenase. *Eur. J. Biochem.* **222 (2):** 345-352.

Cerveñansky, C., and Arias, A. (1984). Glucose-6-phosphate dehydrogenase deficiency in plieotropic carbohydrate-negative mutant strains of *Rhizobium meliloti. J. Bacteriol.* **160 (3):** 1027-1030.

Cevallos, M. A., Encarnación, S., Leija, A., Mora, Y., and Mora, J. (1996). Genetic and Physiological Characterization of a *Rhizobium etli* Mutant Strain Unable to Synthesize Poly-β-Hydroxybutyrate. *J. Bacteriol.* **178 (6):** 1646-1654.

Chen, C.-Y., and Winans, S. C. (1991). Controlled expression of the transcriptional activator gene *virG* in *Agrobacterium tumefaciens* by using the *Escherichia coli lac* promoter. *J. Bacteriol.* **173 (3):** 1139-1144.

Cruz, A. M., Steinmetz, M., Avmerich, S., Richter, R., and le Coq, D. (1990). Induction of levansucrase in *Bacillus subtilis*: an antitermination mechanism negatively controlled by the phosphotransferase system. *J. Bacteriol.* **172 (2):** 1043-1050.

Dassa, E. (1993). Sequence-function relationships in MalG, an inner membrane protein from the maltose transport system in *E. coli. Mol. Microbiol.* **7 (1):** 39-47.

De Jong, A. J., Heidstra, R., Spaink, H. P., Hartog, M. V., Meijer, E. A., Hendricks, T., Schiavo, F. L., Terzi, M., Bisseling, T., Van Kammen, A., and Vries, S. C. (1993). *Rhizobium* lipooligosaccharides rescue a carrot somatic embryo mutant. *Plant Cell* **5 :** 615-620.

de Lorenzo, V., Herrero, M., Jakubzik, U., and Timmis, K. N. (1990). Mini-Tn*5* transposon derivatives for insertional mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol* **172 (11):** 6568-6572.

De Vos, G. F., Walker, G. C., and Signer, E. R. (1986). Genetic manipulations in *Rhizobium meliloti* using two new transposon Tn*5* derivatives. *Mol. Gen. Genet.* **204 :** 485-491.

Debarbouille, M., Arnaud, M., Fouet, A., Klier, A., and Rapoport, G. (1990). The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. *J. Bacteriol.* **172 (7):** 3966-3973.

Debelle, F., Rosenberg, C., Vasse, J., Maillet, F., Martinez, E., Denarie, J., and Truchet, G. (1986). Assignment of symbiotic developmental phenotypes to common and specific nodulation (*nod*) genetic loci of *Rhizobium meliloti. J. Bacteriol.* **168 :** 1075-1086.

Denarie, J., and Cullimore, J. (1993). Lipo-oligosacccharide nodulation factors: A new class of signalling molecules mediating recognition and morphogenesis. *Cell* **74 :** 951-954.

Doherty, D., Leigh, J. A., Glazebrook, J., and Walker, G. C. (1988). *Rhizobium meliloti* mutants that overproduce the *R. meliloti* acidic Calcofluor-binding exopolysaccharide. *J. Bacteriol.* **170 :** 4249-4256.

Dowds, B. C., Sheehan, M. C., Bailey, C. J., and McConnell, D. J. (1988). Cloning and characterization of the gene for a methanol-utilizing alcohol dehydrogenase from *Bacillus stearothermophilus. Gene* **68 (1):** 11-22.

Driscoll, B. T., and Finan, T. M. (1993). NAD^+-dependent malic enzyme of *Rhizobium meliloti* is required for symbiotic nitrogen fixation. *Mol. Micro.* **7 (6):** 865-873.

Dudley, M. E., Jacobs, T. W., and Long, S. R. (1987). Microscopic studies of cell divisions induced in alfalfa roots by *Rhizobium meliloti. Planta* **171 :** 289-301.

Duplay, P., Bedouelle, H., Fowler, A., Zabin, I., Saurin, W., and Hofnung, M. (1984). Sequences of the *malE* gene and of its product, the maltose-binding protein of *Escherichia coli* K12. *J. Biol. Chem.* **259 (16):** 10606-10613.

Ecklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., and Brändén, C.-I. (1976). Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution. *J. Mol. Biol.* **102 :** 27-59.

Egeter, O., and Brueckner, R. (1995). Characterization of a genetic locus essential for maltose-maltotriose utilization in Staphylococcus xylosus. *J. Bacteriol.* **177 (9):** 2408-2415.

Encarnacion, S., Wilms, K., and Mora, J. (1993). Krebs cycle function, poly-3-hydroxybutyrate (PHB) accumulation and oxidative damage in *R. meliloti.* In *New Horizons in Nitrogen Fixation,* R. Palacios, J. Mora and W. E. Newton, eds. (Dordrecht: Kluwer Academic Publishers), pp. 547.

Engelke, T. H., Jagadish, M. N., and Pühler, A. (1987). Biochemical and genetic analysis of *Rhizobium meliloti* mutants defective in C_4-dicarboxylate transport. *J. Gen. Microbiol.* **133 :** 3019-3029.

Erhardt, D. W., Atkinson, E. M., and Long, S. R. (1992). Depolarization of alfalfa root hair membrane potential by *Rhizobium meliloti* Nod factors. *Science* **256 :** 998-1000.

Fath, M. J., and Kolter, R. (1993). ABC transporters: Bacterial exporters. *Micro. Rev.* **57 (4):** 995-1017.

Fennington, G. J., and Hughes, T. A. (1996). The fructokinase from *Rhizobium leguminosarum* biovar *trifolii* belongs to group I fructokinase enzymes and is encoded separately from other carbohydrate metabolism enzymes. *Microbiol.* **142 (2):** 321-330.

Finan, T. M., Hartwieg, E. K., LeMieux, K., Bergman, K., Walker, G. C., and Signer, E. R. (1984). General transduction in *Rhizobium meliloti. J. Bacteriol.* **159 :** 120-124.

Finan, T. M., Hirsch, A. M., Leigh, J. A., Johansen, E., Kuldau, G. A., Deegan, S., Walker, G. C., and Signer, E. R. (1985). Symbiotic mutants of *Rhizobium meliloti* that uncouple plant from bacterial differentiation. *Cell* **40 :** 869-877.

Finan, T. M., Kunkel, B., de Vos, G. F., and Signer, E. R. (1986). Second symbiotic megaplasmid in *Rhizobium meliloti* carrying exopolysaccharide and thiamine synthesis genes. *J. Bacteriol.* **167 :** 66-72.

Finan, T. M., Oresnik, I., and Bottacin, A. (1988). Mutants of *Rhizobium meliloti* defective in succinate metabolism. *J. Bacteriol.* **170 (8):** 3396-3403.

Fisher, R. F., and Long, S. R. (1992). *Rhizobium*-plant signal exchange. *Nature* **357 :** 655-660.

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., Mckenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Geoghagen, N. S. M., Gnehm, C. L., Mcdonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and

Venter, J. C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269 :** 496-498, 507-512.

Franssen, H. J., Vijn, I., Yang, W. C., and Bisseling, T. (1992). Developmental aspects of the *Rhizobium*-legume symbiosis. *Plant Mol. Biol.* **19 :** 89-107.

Freiberg, C., Fellay, R., Bairoch, A., Broughton, W. J., Rosenthal, A., and Perret, X. (1997). Molecular basis of symbiosis between *Rhizobium* and legumes. *Nature* **387 :** 394-401.

Friedman, A. M., Long, S. R., Brown, S. E., Biukema, W. J., and Ausubel, F. M. (1982). Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* **18 :** 289-296.

Gage, D., and Long, S. Personal communication.

Garcia, J. L. (1985). Cloning in *Escherichia coli* and molecular analysis of the sucrose system of the *Salmonella* plasmid SCR-53. *Mol. Gen. Genet.* **201 (3):** 575-577.

Genetics Computer Group. (1991). Program Manual for the GCG Package (Madison, WI: Genetics Computer Group).

Gerngross, T. U., Snell, K. D., Peoples, O. P., Sinskey, A. J., Csuhai, E., Masamune, S., and Stubbe, J. (1994). Overexpression and purification of the soluble polyhydroxyalkanoate synthase from *Alcaligenes eutrophus*: evidence for a required posttranslational modification for catalytic activity. *Biochem.* **33 :** 9311-9320.

Giaquinta, R. T. (1983). Phloem loading of sucrose. *Ann. Rev. Plant Physiol.* **34 :** 347-387.

Gilson, E., Saurin, W., Perrin, D., Bachellier, S., and Hofnung, M. (1991). Palindromic units are part of a new bacterial interspersed mosaic element. *Nucleic Acids Res.* **19 :** 1375-1383.

Gish, W., and States, D. J. (1993). Identification of protein coding regions by database similarity search. *Nature Genet.* **3 :** 266-272.

Glazebrook, J., Ichige, A., and Walker, G. C. (1993). A *Rhizobium melioti* homolog of the *Escherichia coli* peptide-antibiotic transport protein SbmA is essential for bacteroid development. *Genes & Devel.* **7 :** 1485-1497.

Glazebrook, J., Ichige, A., and Walker, G. C. (1996). Genetic analysis of *Rhizobium meliloti bacA-phoA* fusion results in identification of *degP*: two loci required for symbiosis are closely linked to *degP*. *J. Bacteriol.* **178 (3):** 745-752.

Glazebrook, J., and Walker, G. C. (1989). A novel exopolysaccharide can function in place of the Calcofluor-binding exopolysaccharide in nodulation of alfalfa by *Rhizobium meliloti*. *Cell* **56 :** 661-672.

Glazebrook, J., and Walker, G. C. (1991). Genetic techniques in *Rhizobium meliloti*. *Methods Enzymol.* **204 :** 398-418.

Glenn, A. R., and Dilworth, M. J. (1981). The uptake and hydrolysis of disaccharides by fast- and slow-growing species of *Rhizobium*. *Arch. Microbiol.* **129 :** 233-239.

Glucksmann, M. A., Reuber, T. L., and Walker, G. C. (1993a). Family of glycosyl transferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. *J. Bacteriol.* **175 :** 7033-7044.

Glucksmann, M. A., Reuber, T. L., and Walker, G. C. (1993b). Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: A model for succinoglycan biosynthesis. *J. Bacteriol.* **175 :** 7045-7055.

González, J. E., Reuhs, B. L., and Walker, G. C. (1996). Low molecular weight EPS II of *Rhizobium meliloti* allows nodule invasion in *Medicago sativa*. *Proc. Natl. Acad. Sci. USA* **93 (16):** 8636-8641.

Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280 (2):** 309-16.

Henrissat, B., and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293 (3):** 781-8.

Henrissat, B., and Romeu, A. (1995). Families, superfamilies and subfamilies of glycosyl hydrolases. *Biochem. J.* **311 :** 350-351.

Hiles, I. D., Gallagher, M. P., Jamieson, D. J., and Higgins, C. F. (1987). Molecular characterization of the oligopeptide permease of *Salmonella typhimurium*. *J. Mol. Biol.* **195 (1):** 125-142.

Hirosawa, M., Kaneko, T., and Tabata, S. (1995). Cyanobase: Visual presentation of information on the genome of Cyanobacterium *Synechocystis* sp. strain PCC6803 through WWW. In *Proceedings of Genome Informatics Workshop 1995*, pp. 102-103.

Hirsch, A. M. (1992). Developmental biology of legume nodulation. *New. Phytol.* **122 :** 211-237.

Hirsch, A. M., Long, S. R., Bang, M., Haskins, N., and Ausubel, F. M. (1982). Structural studies of alfalfa roots infected with nodulation mutants of *Rhizobium meliloti*. *J. Bacteriol.* **151 :** 411-419.

Hoelzle, I., and Streeter, J. G. (1990). Stimulation of α-glucosidases from fast-growing rhizobia and *Agrobacterium tumefaciens* by K^+, NH_4^+ and Rb^+. *Can. J. Microbiol.* **36 :** 223-227.

Honma, M. A., Asomaning, M., and Ausubel, F. M. (1990). *Rhizobium meliloti nodD* genes mediate host-specific activation of *nodABC*. *J. Bacteriol* **172 :** 901-911.

Hor, L. I., and Shuman, H. A. (1993). Genetic analysis of periplasmic binding protein dependent transport in *Escherichia coli*. Each lobe of maltose-binding protein interacts with a different subunit of the $MalFGK_2$ membrane transport complex. *J. Mol. Biol.* **233 (4):** 659-670.

Houman, F., Diaz-Torres, M. R., and Wright, A. (1990). Transcriptional antitermination in the *bgl* operon of *E. coli* is modulated by a specific RNA binding protein. *Cell* **62 :** 1153-1164.

Hueck, C., Kraus, A., and Hillen, W. (1994). Sequences of *ccpA* and two downstream *Bacillus megaterium* genes with homology to the *motAB* operon from *Bacillus subtilis*. *Gene* **143 (1):** 147-148.

Huisman, G. W., Wonink, E., Meima, R., Kazemier, B., Terpstra, P., and Witholt, B. (1991). Metabolism of poly(3-hydroxyalkanoates) (PHAs) by *Pseudomonas oleovorans*. *J. Biol. Chem.* **266 (4):** 2191-2198.

Ichige, A., and Walker, G. C. (1997). Genetic Analysis of the *Rhizobium meliloti bacA* Gene: Functional Interchangeability with the *Escherichia coli sbmA* Gene and Phenotype of Mutants. *J. Bacteriol.* **179 (1):** 209-216.

Irigoyen, J. J., Emerich, D. W., and Sánchez-Diáz, M. (1992). Phosphoenolpyruvate carboxylase, malate and alcohol dehydrogenase activities in alfalfa (*Medicago sativa*) nodules under water stress. *Physiol. Plant.* **84 :** 61-66.

Irigoyen, J. J., Sanchez-Diaz, M., and Emerich, D. W. (1990). Carbon metabolism enzymes of *Rhizobium meliloti* cultures and bacteroids and their distribution within alfalfa nodules. *Appl. Environ. Microbiol.* **56 (8):** 2587-2589.

Janecek, S., Svensson, B., and Henrissat, B. (1997). Domain evolution in the alpha-amylase family. *J. Mol. Evolution* **45 (3):** 322-331.

Jelesko, J. G., and Leigh, J. A. (1994). Genetic characterization of a *Rhizobium meliloti* lactose utilization locus. *Mol. Micro.* **11 (1):** 165-173.

Jendrossek, D., Steinbüchel, A., and Schlegel, H. G. (1988). Alcohol dehydrogenase gene from *Alcaligenes eutrophus*: subcloning, heterologous expression in *Escherichia coli*, sequencing and location of Tn*5* insertions. *J. Bacteriol.* **170 (11):** 5248-5256.

Jording, D., Uhde, C., Schmidt, R., and Pühler, A. (1994). The C_4-dicarboxylate transport system of *Rhizobium meliloti* and its role in nitrogen fixation during symbiosis with alfalfa (*Medicago sativa*). *Experientia* **50 :** 874-883.

Joshi, P. A., Caetano-Anollés, G., Graham, E. T., and Gresshoff, P. M. (1991). Ontogeny and ultrastructure of spontaneous nodules in alfalfa (*Medicago sativa*). *Protoplasma* **162 :** 1-11.

Kahn, D., Batut, J., Daveran, M. L., and Fourment, J. (1993). Structure and regulation of the *fixNOQP* operon from *Rhizobium meliloti*. In *New Horizons in Nitrogen Fixation*, R. Palacios, J. Mora and W. E. Newton, eds. (Dordrecht: Kluwer Academic Publishers), p. 474.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3 (3):** 109-136.

Kannan, R., Mukundan, G., Ait-Abdelkader, N., Augier-Magro, V., Baratti, J., and Gunasekaran, P. (1995). Molecular cloning and characterization of the extracellular sucrase gene (*sacC*) of *Zymomonas mobilis*. *Arch. Microbiol.* **163 (3):** 195-204.

Kawamukai, M., Utsumi, R., Takeda, K., Higashi, A., Matsuda, H., Choi, Y. L., and Komano, T. (1991). Nucleotide sequence and characterization of the *sfs1* gene: *sfs1* is inovolved in CRP*-dependent *mal* gene expression in *Escherichia coli*. *J. Bacteriol.* **173 (8):** 2644-2648.

Kellerman, and Ferenci (1982). Maltose binding protein from *E. coli*. *Methods Enzymol.* **90 :** 459-463.

Kim, K.-S., and Farrand, S. K. (1996). Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by *Agrobacterium tumefaciens* are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor. *J. Bacteriol.* **178 (11):** 3275-3284.

Kizaki, H., Hata, Y., Watanabe, K., Katsube, K., and Suzuki, Y. (1993). Polypeptide folding of *Bacillus cereus* ATCC7064 oligo-1,6-glucosidase revealed by 3.0 Å resolution X-ray analysis. *J. Biochem.* **113 :** 646-649.

Law, J. H., and Slepeckey, R. A. (1961). Assay of poly-β-hydroxybutyric acid. *J. Bacteriol.* **82 :** 33-36.

le Rudulier, D. Personal communication.

Legocki, R. P., and Verma, D. P. S. (1980). Identification of nodule specific host proteins (nodulins) involved in the development of *Rhizobium*-legume symbiosis. *Cell* **20 :** 153-163.

Leigh, J., and Walker, G. C. Unpublished results.

Leigh, J. A., and Coplin, D. C. (1992). Exopolysaccharides in plant-bacterial interactions. *Annu. Rev. Microbiol.* **46 :** 307-346.

Leigh, J. A., Reed, J. W., Hanks, J. F., Hirsch, A. M., and Walker, G. C. (1987). *Rhizobium meliloti* mutants that fail to succinylate their Calcofluor-binding exopolysaccharide are defective in nodule invasion. *Cell* **51 :** 579-587.

Leigh, J. A., Signer, E. R., and Walker, G. C. (1985). Exopolysaccharide-deficient mutants of *Rhizobium meliloti* that form ineffective nodules. *Proc. Nat. Acad. Sci. USA* **82 :** 6231-6235.

Lemoigne, M. (1925). Etudes sur l'autolyse microbienne acidification par formation d'acide b-oxybutyrique. *Ann. Inst. Pasteur Paris* **39 :** 144.

Lemoigne, M. (1926). Products of dehydration and of polymerization of β-hydroxybutyric acid. *Bull. Soc. Chem. Biol.* **8 :** 770-782.

Lengeler, J. W., Mayer, R. J., and Schmid, K. (1982). Phosphoenolpyruvate-dependent phosphotransferase system enzyme III and plasmid-encoded sucrose transport in *Escherichia coli. J. Bacteriol* **151 (468-471):**

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., and Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **344 :** 781-784.

Long, S., Reed, J. W., Himawan, J., and Walker, G. C. (1988). Genetic analysis of a cluster of genes required for synthesis of the Calcofluor-binding exopolysaccharide of *Rhizobium meliloti. J. Bacteriol.* **170 :** 4239-4248.

Lullien, V., Barker, D. G., de Lajudie, P., and Huguet, T. (1987). Plant gene expression in effective and ineffective root nodules of alfalfa (*Medicago sativa*). *Plant Mol. Biol.* **9 :** 469-478.

Manchenko, G. P. (1994). *Handbook of detection of enzymes on electrophoretic gels* (Boca Raton: CRC Press).

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). *Molecular cloning: a laboratory manual* (Cold Spring Harbor, N. Y.: Cold Spring Harbor Laboratory).

Manoil, C., and Beckwith, J. (1985). Tn*phoA*: A transposon probe for protein export signals. *Proc. Natl. Acad. Sci. USA* **82 :** 8129-8133.

Margolin, W., and Long, S. R. (1994). *Rhizobium meliloti* contains a novel second homolog of the cell division gene *ftsZ. J. Bacteriol.* **176 :** 2033-2043.

Martinez-de Drets, G., Arias, A., and de Cutinella, M. R. (1974). Fast- and slow-growing rhizobia: differences in sucrose utilization and invertase activity. *Can. J. Microbiol.* **20 :** 605- 609.

Meade, H. M., Long, S. R., Ruvkun, G. B., Brown, S. E., and Ausubel, F. M. (1982). Physical and genetic characterization of symbiotic and auxotrophic mutants of *Rhizobium meliloti* induced by transposon Tn*5* mutagenesis. *J. Bacteriol.* **149 :** 114-122.

Mergaert, P., Van Montagu, M., Promé, J. C., and Holsters, M. (1993). Three unusual modifications, a D-arabinosyl, an N-methyl, and a carbamoyl group, are present on the Nod factors of *Azorhizobium caulinodans* strain ORS571. *Proc. Natl. Acad. Sci. USA* **90 (1551-1555):**

Meyer, R. J., and Shapiro, J. A. (1980). Genetic organization of the broad-host-range IncP-1 plasmid R751. *J. Bacteriol.* **143 (3):** 1362-1373.

Mikkat, S., Effmert, U., and Hagemann, M. (1997). Uptake and use of the osmoprotective compounds trehalose, glucosylglycerol, and sucrose by the cyanobacterium *Synechocystis* sp. PCC6803. *Arch. Microbiol.* **167 :** 112-118.

Nakao, M., Nakayama, T., Kakudo, A., Inohara, M., Harada, M., Omura, F., and Shibano, Y. (1994). Structure and expression of a gene coding for thermostable alpha-glucosidase with a broad substrate specificity from *Bacillus* sp. SAM1606. *Eur. J. Biochem.* **220 (2):** 293-300.

Nikaido, H. (1994). Maltose transport system of *Escherichia coli*: an ABC-type transporter. *FEBS Lett* **346 (1):** 55-8.

Norris, J. H., Macol, L. A., and Hirsch, A. M. (1988). Nodulin gene expression in effective alfalfa nodules and in nodules arrested at three different stages of development. *Plant Physiol.* **88 :** 321-328.

O'Brian, M. R. (1996). Heme synthesis in the *Rhizobium*-legume symbiosis: a palette for bacterial and eukaryotic pigments. *J. Bacteriol.* **178 (9):** 2471-2478.

Østerås, M., Stanley, J., and Finan, T. M. (1995). Identification of rhizobium-specific intergenic mosaic elements (RIME) within an essential two-component regulatory system of *Rhizobium* species. *J. Bacteriol.* **177 (19):** 5485-5494.

Ozols, J. (1993). Isolation and the complete amino acid sequence of lumenal endoplasmic reticulum glucose-6-phosphate dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **90 (11):** 5302-5306.

Paau, A. S., Cowles, J. R., and Raveed, D. (1978). Development of bacteroids in alfalfa *(Medicago sativa)* nodules. *Plant Physiol.* **62 :** 526-530.

Palchaudhuri, S., Rahn, S., Santos, D. S., and Maas, W. K. (1977). Characterization of plasmids in a sucrose-fermenting strain of *Escherichia coli*. *J. Bacteriol.* **130 :** 1402-1403.

Peoples, O. P., Masamune, S., Walsh, C. T., and Sinskey, A. J. (1987). Biosynthetic thiolase from *Zooglea ramigera*. III. Isolation and characterization of the structural gene. *J. Biol. Chem.* **262 :** 97-102.

Peoples, O. P., and Sinskey, A. J. (1989a). Fine structural analysis of the *Zooglea ramigera phbA-phbB* locus encoding β-ketothiolase and acetoacetyl CoA reductase: nucleotide sequence of *phbB*. *Mol. Microbiol.* **3 (3):** 349-357.

Peoples, O. P., and Sinskey, A. J. (1989b). Poly-β-hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus* H16: Identification and characterization of the PHB polymerase gene (*phbC*). *J. Biol. Chem.* **264 (26):** 15298-15303.

Peoples, O. P., and Sinskey, A. J. (1989c). Poly-β-hydroxybutyrate biosynthesis in *Alcaligenes eutrophus* H16: Characterization of the genes encoding β-ketothiolase and acetoacetyl-CoA reductase. *J. Biol. Chem.* **264 (26):** 15293-15297.

Peters, N. K., Frost, J. W., and Long, S. R. (1986). A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* **233 :** 977-980.

Petrovics, G., Putnoky, P., Reuhs, B., Kim, J., Thorp, T. A., Noel, K. D., Carlson, R. W., and Kondorosi, A. (1993). A novel type of surface polysaccharide involved in symbiotic nodule development requires the expression of new fatty acid synthase-like gene cluster in *Rhizobium meliloti*. *Mol. Microbiol.* **8 :** 1083-1094.

Philipp, W. J., Poulet, S., Eiglmeier, K., Pascopella, L., Balasubramanian, V., Heym, B., Bergh, S., Bloom, B. R., Jacobs, W. R. J., and Cole, S. T. (1996). An integrated map of the genome of the tubercle bacillus, *Mycobacterium tuberculosis* H37Rv, and comparison with *Mycobacterium leprae*. *Proc. Natl. Acad. Sci. U.S.A.* **93 (7):** 3132-3137.

Pocard, J. A., Vincent, N., Boucompagni, E., Tombras-Smith, L., Poggi, M. C., and le Rudulier, D. (1997). Molecular characterization of the *bet* genes encoding glycine betaine synthesis in *Sinorhizobium meliloti* 102F34. *Microbiol.* **143 (4):** 1369-1379.

Postma, P. W., and Lengeler, J. W. (1985). Phosphoenolpyruvate carbohydrate phosphotransferase system of bacteria. *Microbiol. Rev.* **49 :** 232-269.

Postma, P. W., Lengeler, J. W., and Jacobson, G. R. (1993). Phosphoenol-pyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57 (3):** 543-594.

Poulsen, P., and Kreiberg, J. D. (1993). Starch branching enzyme cDNA from *Solanum tuberosum*. *Plant Physiol.* **102 :** 1053-1054.

Povolo, S., Tombolini, R., Morea, A., Anderson, A. J., Casella, S., and Nuti, M. P. (1994). Isolation and characterization of mutants of *Rhizobium meliloti* unable to synthesize poly-β-hydroxybutyrate. *Can. J. Microbiol.* **40 :** 823-829.

Putnoky, P., Grosskopf, E., Cam-Ha, D. T., Kiss, G. B., and Kondorosi, A. (1988). *Rhizobium fix* genes mediate at least two communication steps in symbiotic nodule development. *J. Cell Biol.* **106 :** 597-607.

Quiocho, F. A., and Ledvina, P. S. (1996). Atomic structure and specificity of bacterial periplasmic receptors for active transport and chemotaxis: variation of common themes. *Mol. Microbiol.* **20 (1):** 17-25.

Rastogi, V. K., and Watson, R. J. (1991). Aspartate aminotransferase activity is required for aspartate catabolism and symbiotic nitrogen fixation in *Rhizobium meliloti. J. Bacteriol.* **173 :** 2879-2887.

Redmond, J. W., Batley, M., Djordjevic, M. A., Innes, R. W., Kuempel, P. L., and Rolfe, B. G. (1986). Flavones induce expression of nodulation genes in *Rhizobium. Nature* **323 :** 632-635.

Reed, J. W., and Walker, G. C. (1991a). Acidic conditions permit effective nodulation of alfalfa by invasion-deficient *Rhizobium meliloti exoD* mutants. *Genes Dev.* **5 :** 2274-2287.

Reed, J. W., and Walker, G. C. (1991b). The *exoD* gene of *Rhizobium meliloti* encodes a novel function needed for alfalfa nodule invasion. *J. Bacteriol.* **1991 :** 664-677.

Reibach, P. H., and Streeter, J. G. (1983). Metabolism of ^{14}C-labeled photosynthate and distribution of enzymes of glucose metabolism in soybean nodules. *Plant Physiol.* **72 :** 634.

Reid, M. F., and Fewson, C. A. (1994). Molecular characterization of microbial alcohol dehydrogenases. *Crit. Rev. Microbiol.* **20 (1):** 13-56.

Reizer, J., Saier, M. H. J., Deutscher, J., Grenier, F., Thompson, J., and Hengstenberg, W. (1988). The phosphoenolpyruvate sugar phosphotransferase system in Gram-positive bacteria: properties, mechanism and regulation. *CRC Crit. Rev. Biochem.* **15 :** 297-338.

Rigaud, J., and Trinchant, J.-C. (1973). Isolement d'une alcool déshydrogénase chez *Rhizobium* et rôle dans le métabolisme indolique. *Physiol. Plant.* **28 :** 160-165.

Robinson, G. A., Bailey, C. J., and Dowds, B. C. (1994). Gene structure and amino acid sequences of alcohol dehydrogenases of *Bacillus stearothermophilus. Biochim. Biophys. Acta* **1218 (3):** 432-434.

Romanov, V. I., Hajy-zadeh, B. R., Ivanov, B. F., Shaposhnikov, G. L., and Kretovich, W. L. (1985). Labelling of lupine nodule metabolites with $^{14}CO_2$ assimilated from the leaves. *Phyto.* **24 (10):** 2157-2160.

Romanov, V. I., Yushkova, L. A., and Kretovich, V. L. (1974). Relationship between the content of poly-β-hydroxybutyric acid in the bacteroids of *Rhizobium lupini,*

their respiration, and nitrogen fixation. *Doklady Akademii Nauk SSSR* **216 (3):** 694-697.

Ronson, C. W., Lyttleton, P., and Robertson, J. G. (1981). C_4-dicarboxylate transport mutants of *Rhizobium trifolii* form ineffective nodules on *Trifolium repens*. *Proc. Nat. Acad. Sci. USA* **82 :** 6231-6245.

Ronson, C. W., and Primrose, S. B. (1979). Carbohydrate metabolism in *Rhizobium trifolii*: identification and symbiotic properties of mutants. *J. Gen. Microbiol.* **112 :** 77-88.

Rosenberg, C., and Huguet, T. (1984). The pAtC58 plasmid is not essential for tumor induction. *Mol. Gen. Genet.* **196 :** 533-536.

Russell, P. R., and Hall, B. D. (1983). The primary structure of the alcohol dehydrogenase gene from the fission yeast *Schizosaccharomyces pombe*. *J. Biol. Chem.* **258 :** 143-149.

Russell, R. R., Aduse-Opoku, J., Sutcliffe, I. C., Tao, L., and Ferretti, J. J. (1992). A binding protein-dependent transport system in *Streptococcus mutans* responsible for multiple sugar metabolism. *J. Biol. Chem.* **267 (7):** 4631-4637.

Saier, M. H., Jr., and Reizer, J. (1994). The bacterial phosphotransferase system: new frontiers 30 years later. *Mol. Micro.* **13 (5):** 755-764.

Sakoda, H., and Imanaka, T. (1992). Cloning and sequencing of the gene coding for alcohol dehydrogenase of *Bacillus stearothermophilus* and rational shift of the optimum pH. *J. Bacteriol.* **174 (4):** 1397-1402.

Sanjuan, J., Carlson, R. W., Spaink, H. P., Bhat, U. R., Barbour, W. M., Glushka, J., and Stacey, G. (1992). A 2-*O*-methylfucose moiety is present in the lipo-oligosaccharide nodulation signal of *Bradyrhizobium japonicum*. *Proc. Nat. Acad. Sci. USA* **89 :** 8789-8793.

Saurin, W., and Dassa, E. (1994). Sequence relationships between integral inner membrane proteins of binding protein-dependent transport systems: evolution by recurrent gene duplications. *Protein Sci.* **3 (2):** 325-344.

Scharff, A. J., Rodseth, L. E., Spurlino, J. C., and Quiocho, F. A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry* **31 (44):** 10657-10663.

Schlegel, H. G., and Steinbuchel, A. (1991). Physiology and molecular genetics of poly(β-hydroxy-alkanoic acid) synthesis in *Alcaligenes eutrophus*. *Molec. Microbiol.* **80 :** 161-169.

Schmid, K., Schupfner, M., and Schmitt, R. (1982). Plasmid-mediated uptake and metabolism of sucrose by *Eschericia coli* K-12. *J. Bacteriol.* **151 (1):** 68-76.

Scholle, R. R., Robb, S. M., Robb, F. T., and Woods, D. R. (1989). Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrase gene (*scrB*). *Gene* **80 (1):** 49-56.

Schubert, P., Steinbüchel, A., and Schlegel, H. G. (1988). Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-β-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. *J. Bacteriol.* **170 (12):** 5837-5847.

Schuerman, P. L., Liu, J. S., Mou, H., and Dandekar, A. M. (1997). 3-ketoglycoside-mediated metabolism of sucrose in *E. coli* as conferred by genes from *Agrobacterium tumefaciens*. *Appl. Microbiol. Biotechnol.* **47 (5):** 560-565.

Schwedock, J. S., and Long, S. R. (1990). ATP sulphurylase activity of the *nodP* and *nodQ* gene products of *Rhizobium meliloti*. *Nature* **348 :** 644-647.

Semino, C. E., and Robbins, P. W. (1995). Synthesis of "Nod"-like chitin oligosaccharides by the *Xenopus* developmental protein DG42. *Proc. Natl. Acad. Sci. USA* **92 (8):** 3498-3501.

Semino, C. E., Specht, C. A., Raimondi, A., and Robbins, P. W. (1996). Homologs of the *Xenopus* developmental gene DG42 are present in zebrafish and mouse and are involved in the synthesis of Nod-like chitin oligosaccharides during early embryogenesis. *Proc. Natl. Acad. Sci. USA* **93 (10):** 4548-4553.

Shen, W.-C., Stanford, D. R., and Hopper, A. K. (1996). Los1p, involved in yeast pre-tRNA splicing, positively regulates members of the SOL gene family. *Genetics* **143 (699-712):**

Shuman, H. A. (1982). Active transport of maltose in *Escherichia coli* K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. *J. Biol. Chem.* **257 (10):** 5455-5461.

Smith, L. T., Smith, G. M., D'Souza, M. R., Pocard, J.-A., le Rudulier, D., and Madkour, M. A. (1994). Osmoregulation in *Rhizobium meliloti*: Mechanism and control by other environmental signals. *J. Exp. Zool.* **268 :** 162-165.

Sobral, B. W. S., Honeycutt, R. J., Atherly, A. G., and McClelland, M. (1991). Electrophoretic separation of the three *Rhizobium meliloti* replicons. *J. Bacteriol.* **173 (16):** 5173-5180.

Spaink, H. P., Bloemberg, G. V., van Brussel, A. A. N., Lugtenberg, B. J. J., van der Drift, K. M. G. M., Haverkamp, J., and Thomas-Oates, J. E. (1995). Host specificity of *Rhizobium leguminosarum* is determined by the hydrophobicity of highly unsaturated fatty acyl moieties of the nodulation factors. *Mol. Plant-Microbe Int.* **8 :** 155-164.

Spaink, H. P., Sheeley, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J. (1991). A novel

highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature* **354 :** 125-129.

Spurlino, J. C., Lu, G. Y., and Quiocho, F. A. (1991). The 2.3-Å resolution structure of the maltose- or maltodextrin binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266 (8):** 5202-5219.

Steinbüchel, A., Hustede, E., Liebergesell, M., Pieper, U., Timm, A., and Valentin, H. (1992). Molecular basis for biosynthesis and accumulation of polyhydroxyalkanoic acids in bacteria. *FEMS Micro. Rev.* **103 :** 217-230.

Steinbüchel, A., and Schlegel, H. G. (1989). Excretion of pyruvate by mutants of *Alcaligenes eutrophus*, which are impaired in the accumulation of poly(β-hydroxybutyric acid) (PHB), under conditions permitting syntheis of PHB. *Appl. Microbiol. Biotechnol.* **31 :** 168-175.

Stowers, M. D. (1985). Carbon metabolism in *Rhizobium* species. *Ann. Rev. Microbiol.* **39 :** 89-108.

Streeter, J. G. (1991). Transport and metabolism of carbon and nitrogen in legume nodules. In *Advances in Botanical Research*, J. A. Callow, ed. (London: Academic Press), pp. 129-187.

Sung, M. H., Tanizawa, K., Tanaka, H., Kuramitsu, S., Kagamiyama, H., Hirotsu, K., Okamoto, A., Higuchi, T., and Soda, K. (1991). Thermostable aspartate aminotransferease from a thermophilic *Bacillus* species. Gene cloning, sequence determination, and preliminary X-ray characterization. *J. Biol. Chem.* **266 :** 2567-2572.

Suzuki, Y., Ito, N., Yuuku, T., Yamagata, H., and Udaka, S. (1989). Amino acid stabilizing a *Bacillus* α-amylase against irreversible thermoinactivation. *J. Biol. Chem.* **264 :** 18933-18938.

Svensson, B. (1994). Protein engineering in the α-amylase family:catalytic mechanism, substrate specificity, and stability. *Plant Mol. Biol.* **25 :** 141-157.

Szmelcman, S., Sassoon, N., and Hofnung, M. (1997). Residues in the alpha helix 7 of the bacterial maltose binding protein which are important in interactions with the Mal FGK2 complex. *Protein Sci.* **6 (3):** 628-636.

Takii, Y., Takahashi, K., Yamamoto, K., Sogabe, Y., and Suzuki, Y. (1996). *Bacillus stearothermophilus* ATCC12016 alpha-glucosidase specific for alpha-1,4 bonds of maltosaccharides and alpha-glucans shows high amino acid sequence similarities to seven alpha-D-glucohydrolases with different substrate specificity. *Appl. Microbiol. Biotechnol.* **44 :** 629-634.

Timm, A., and Steinbüchel, A. (1992). Cloning and molecular analysis of the poly(3-hydroxyalkanoic acid) gene locus of *Pseudomonas aeruginosa* PAO1. *Eur. J. Biochem.* **209 :** 15-30.

Titgemeyer, F., Jahreis, K., Ebner, R., and Lengeler, J. W. (1996). Molecular analysis of the scrA and scrB genes from Klebsiella pneumoniaw and plasmid pUR400 which encode the sycrose transport protein Enzyme II(Scr) of the phosphotransferase system and a sucrose-6-phosphate invertase. *Mol. Gen. Genet.* **250 (2):** 197-206.

Tombolini, R., and Nuti, M. P. (1989). Poly(β-hydroxyalkanoate) biosynthesis and accumulation by different *Rhizobium* species. *FEMS Microbiol. Lett.* **60 :** 299-304.

Tombolini, R., Povolo, S., Buson, A., Squartini, A., and Nuti, M. P. (1995). Poly-β-hydroxybutyrate (PHB) biosynthetic genes in *Rhizobium meliloti* 41. *Microbiology* **141 :** 2553-2559.

Truchet, G., Debellé, F., Vasse, J., Terzaghi, B., Garnerone, A.-M., Rosenberg, C., Batut, J., Maillet, F., and Dénarié, J. (1985). Identification of a *Rhizobium meliloti* pSym2011 region controlling the host specificity of root hair curling and nodulation. *J. Bacteriol.* **164 (3):** 1200-1210.

Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., de Billy, F., Promé, J.-C., and Dénarié, J. (1991). Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351 :** 670-673.

Ucker, D. S., and Signer, E. R. (1978). Catabolite-repression-like phenomenon in *Rhizobium meliloti*. *J. Bacteriol.* **136 :** 1197-1200.

Valentin-Hansen, P., Larsen, J. E., Hojrup, P., Short, S. A., and Barbier, C. S. (1986). Nucleotide sequence of the CytR regulatory gene of *E. coli* K-12. *Nucleic Acids Res.* **14 (5):** 2215-2228.

van Rhijn, P., and Vanderleyden, J. (1995). The *Rhizobium*-Plant Symbiosis. *Microbiological Reviews* **59 (1):** 124-142.

Vasse, J., deBilly, F., Camut, S., and Truchet, G. (1990). Correlation between ultrastructural differentiation of bacteroids and nitrogen fixation in alfalfa nodules. *J. Bacteriol.* **172 :** 4295-4306.

Verma, D. P. S., Haugland, R., Brisson, N., Legocki, R. P., and Lacroix, L. (1981). Regulation of the expression of leghemoglobin genes in effective and ineffective root nodules. *Biochim. Biophys. Acta* **653 :** 98-107.

Vijn, I., Neves, L. d., Kammen, A. v., Franssen, H., and Bisseling, T. (1993). Nod factors and nodulation in plants. *Science* **260 :** 1764-1765.

Watanabe, K., Kitamura, K., and Suzuki, Y. (1996). Analysis of the critical sites for protein thermostabilization by proline substitution in oligo-1,6-glucosidase from *Bacillus coagulans* ATCC 7050 and the evolutionary consideration of proline residues. *Appl. Environ. Microbiol.* **62 (6):** 2066-2073.

Williams, S. G., Greenwood, J. A., and Jones, C. W. (1992). Molecular analysis of the *lac* operon encoding the binding-protein-dependent lactose transport system and β-galactosidase in *Agrobacterium radiobacter*. *Mol. Microbiol.* **6 (13):** 1755-1768.

Williams, S. G., Greenwood, J. A., and Jones, C. W. (1995). *Agrobacterium radiobacter* and related organisms take up fructose via a binding-protein-dependent active-transport system. *Microbiol.* **141 (10):** 2601-2610.

Wright, A. Personal communication.

Xu, K., He, Z. Q., Mao, Y. M., Sheng, R. Q., and Sheng, Z. J. (1993). On two transposable elements from *Bacillus stearothermophilus*. *Plasmid* **29 (1):** 1-9.

Yang, C., Signer, E. R., and Hirsch, A. M. (1992). Nodules initiated by *Rhizobium meliloti* exopolysaccharide mutants lack a discrete, persistent nodule meristem. *Plant Physiol.* **98 :** 143-151.

Yarosch, O. K., Charles, T. C., and Finan, T. M. (1989). Analysis of C_4-dicarboxylate transport genes in *Rhizobium meliloti*. *Mol. Microbiol.* **3 :** 813-823.

York, G. M., and Walker, G. C. (1997). The *Rhizobium meliloti exoK* and *prsD/prsE/exsH* genes are components of independent degradative pathways which contribute to production of low-molecular-weight succinoglycan. *Mol. Microbiol.* **25 :** 117-134.

Zhang, L., Radziejewska-Lebrecht, J., Krajewska-Pietrasik, D., Toivanen, P., and Skurnik, M. (1997). Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O:8. *Mol. Microbiol.* **23 (1):** 63-76.